А.Н.Гусев

Дисперсионный анализ в экспериментальной психологии

Учебное пособие

Для студентов факультетов психологии высших учебных заведений по направлению 521000 — «Психология»

Москва
Учебно-методический коллектор «Психология»
2000

УДК 159.938
ББК 88

Рекомендовано кафедрой психологии личности факультета психологии Московского государственного университета имени М.В.Ломоносова.

Рецензенты:
А.П.Кулаичев, кандидат физико-математических наук;
А.Н.Кричевец, доктор философских наук, кандидат физико-математических наук ;
А.М.Черноризов, доктор психологических наук

Гусев А.Н.
Г962 Дисперсионный анализ в экспериментальной психологии: Учебное пособие для студентов факультетов психологии высших учебных заведений по направлению 521000 — «Психология» . — М.: Учебно-методический коллектор «Психология», 2000. — 136 с.

ISBN 5-93692-015-1

В учебном пособии представлены современные методы дисперсионного анализа как мощного средства оценки факторных эффектов в психологических экспериментах. Кроме общего описания различных процедур дисперсионного анализа даны подробные рекомендации по использованию двух наиболее популярных среди отечественных психологов статистических систем — Stadia и SPSS. Приведены примеры обработки данных психологических исследований методами однофакторного, многофакторного и многомерного дисперсионного анализа, показана специфика обработки данных факторных экспериментов с повторными измерениями.

Рекомендуется студентам-психологам в качестве учебного пособия по курсу «Математические методы в психологии», а также для самостоятельного изучения статистических методов обработки данных и анализа результатов. Кроме того, книга будет полезна психологам, проводящим эмпирические исследования, в которых оцениваются воздействия на психологические переменные различного рода факторных эффектов.

ISBN 5-93692-015-1

Есть только один способ достижения
счастья на этом земном шаре —
иметь ясное сознание
или не иметь его вовсе.

Огден Нэш

Предисловие

Дисперсионный анализ (ДА) является мощным современным статистическим методом обработки и анализа экспериментальных данных в психологии, биологии, медицине и других науках. Он очень тесно связан с конкретной методологией планирования и проведения экспериментальных исследований. Вместе с тем, наш опыт общения с коллегами-психологами, студентами и аспирантами факультета психологии Московского государственного университета имени М.В.Ломоносова показал, что многие из них недостаточно уверенно владеют этим методом и поэтому зачастую просто боятся его использовать.

Замысел данного учебного пособия возник у автора случайно и некоторым образом даже непроизвольно. Первоначальной предпосылкой его написания явилась, с одной стороны, большая личная неудовлетворенность той литературой, которая имеется на русском языке по применению дисперсионного анализа в практике научных исследований, в том числе в психологии. С другой стороны, в последнее десятилетие появились современные статистические системы, в которых представлены новые процедуры ДА. Однако, даже по фирменным описаниям этих статистических систем крайне сложно разобраться в специфике и многообразии предлагаемых вариантов обработки данных. У нас и у многих наших коллег возникали постоянные трудности с правильным пониманием многочисленных опций в процедурах ДА, появив-

шихся в различных статистических системах. Более того, целый ряд новых терминов, встречавшихся в их не очень-то подробных и содержательных описаниях, был нам просто незнаком. Попытки обратиться к доступным для нас специалистам-статистикам за разъяснениями не увенчались успехом, поскольку эти достаточно новые статистические процедуры ими были также еще детально не освоены. Особенно это касалось процедуры ДА с повторными измерениями.

В то же время, судя по публикациям результатов экспериментальных исследований в большинстве солидных международных психологических журналов, использование современных процедур ДА стало не только нормой, но и обязательным требованием представления результатов. Кроме того, опыт общения с зарубежными коллегами-преподавателями показал нам, что обучение студентов основам ДА и его использованию с помощью современных компьютерных статистических систем стало обычным делом для многих ведущих университетов. К тому же в последние 5—7 лет появились и новые учебники по статистике для психологов, где нашли отражение самые современные варианты ДА.

Таким образом, у нас появилась потребность систематизировать свои разрозненные записи о различных вариантах ДА в виде учебного пособия, чтобы в какой-то мере компенсировать указанные выше недостатки в подготовке студентов-психологов. Многочисленные консультации по использованию ДА в психологии, которые нам приходится давать и студентам, и своим коллегам-преподавателям, только убедили нас в необходимости издания пособия такого рода.

Безусловно, мы не тешим себя надеждой, что нам удалось в достаточной степени полно и систематично изложить основы данной группы методов и рассмотреть все их нюансы и трудности. Тем не менее, нам представляется, что настоящая книга будет полезна тем, кто хочет разобраться в специфике использования ДА и получить рекомендации по работе с компьютерными статистическими системами. На последнем остановимся чуть подробнее. Мы выбрали две статистические системы, на примере которых показываем, как использовать различные варианты ДА. Это отечественная система

STADIA (профессиональная версия 6.0) и американская система SPSS (русская версия 8.01). Свой выбор мы сделали намеренно, по двум причинам. Во-первых, это наиболее распространенные среди психологов статистические системы для персональных компьютеров. Во-вторых, к ним имеется содержательная и доступная российскому пользователю литература (см.: *Кулаичев*, 1998; *Тюрин*, *Макаров*, 1998; Руководство пользователя. SPSS Base 8.0, 1998).

Именно по причине большей доступности русскоязычной версии SPSS для отечественного пользователя мы приводим примеры обработки экспериментальных данных в версии SPSS 8.01, а не в более новых версиях. Наш собственный опыт работы с SPSS показал, что никаких принципиальных различий в использовании процедур ДА в системе SPSS, начиная с версии 8.0, по сравнению с версиями 9.0 и 10.0, нет. В новых версиях лишь немного изменились пункты основного меню, но это уже, как говорят психофизики, «едва заметные различия», никак не влияющие на результаты вычислений.

Хочется отметить, что работа по модификации отечественной статистической системы STADIA продолжается: в последние дни работы над рукописью нам стало известно, что ее автор А.П.Кулаичев дополнительно включил в набор процедур ДА вариант ДА с повторными измерениями и начал работу над алгоритмом многомерного ДА. Поэтому мы надеемся, что к моменту выхода данного учебного пособия психологи смогут получить доработанную версию STADIA.

В заключение нам хотелось бы выразить искреннюю благодарность своим коллегам, которые явно или неявно способствовали работе над книгой. Это прежде всего С.А.Шапкин, с которым нам пришлось не раз «продираться» через премудрости опций статистических пакетов и руководств к ним, а также А.П.Кулаичев, которому в течение многих лет приходилось отвечать на множество наших наивных и непрофессиональных вопросов. Кроме того, мы выражаем благодарность руководителю российского представительства фирмы SPSS в Москве Антону Ковтуну за предоставленную возможность работать со всеми версиями системы SPSS и их описаниями.

Большая благодарность И.В.Тимофееву за замечания, сделанные им после прочтения рукописи.

В работе над учебным пособием нам помогла блестящая книга Дж.Хауэлла «Статистические методы для психологов» (*Howell*, 1998).

Особая благодарность нашим рецензентам — кандидату физико-математических наук А.П.Кулаичеву, доктору философских наук, кандидату физико-математических. наук А.Н.Кричевцу и доктору психологических наук А.М.Черноризову за внимательное прочтение книги и сделанные замечания.

Глава 1

Сущность и логика дисперсионного анализа

1.1. Основные понятия

Дисперсионный анализ, или *ANOVA* (аббревиатура английского выражения Analysis of Variance), давно приобрел статус одного из самых используемых методов статистического анализа в психологии. Популярность и полезность этой техники обработки эмпирических данных обусловлены по крайней мере двумя причинами. Во-первых, ANOVA, подобно t-критерию Стьюдента, позволяет оценить различия между выборочными средними; однако, в отличие от t-критерия, в нем *нет ограничений на количество сравниваемых средних.* Таким образом, вместо того, чтобы поставить вопрос о различии двух выборочных средних, мы можем оценить, различаются ли два, три четыре, пять или k средних. Во-вторых, ANOVA позволяет нам иметь дело с двумя или более независимыми переменными (признаками, факторами) одновременно, оценивая не только *эффект каждой из них по отдельности*, но и *эффекты взаимодействия* между ними.

Впервые дисперсионный анализ (ДА) был разработан американским статистиком Р.Фишером (1925) для анализа результатов агротехнических опытов и затем усовершенствован многими его последователями. Как следует из названия метода, его суть состоит в *разложении (анализе) дисперсии* одной или нескольких переменных на составляющие компоненты, сравнивая которые друг с другом с помощью *F-критерия,* можно оценить ее (их) вклад в общую вариацию данных. Поскольку ДА является комплексным методом оценки

различных экспериментальных взаимодействий, он требует определенной группировки данных, из чего следует необходимость специальной организации (планирования) эмпирического исследования. Результаты психологических измерений, подвергаемые ДА, должны быть сгруппированы с учетом выделения определенных контролируемых в исследовании факторов (или возможного взаимодействия нескольких факторов) и количеством повторных наблюдений. Для применения ряда современных вариантов многофакторного ДА бывает необходима реализация достаточно сложных и трудоемких планов проведения эмпирического исследования (см., напр., *Налимов*, 1971; *Kirk*, 1995).

Воспользуемся данными одного вымышленного психологического исследования, чтобы на конкретном примере показать, какую задачу решает ДА и ввести ряд необходимых терминов. Хотя наш пример касается лишь однофакторного ДА (см. следующий параграф), он, тем не менее, позволит описать ряд инвариантных понятий, важных для всего последующего изложения. В эксперименте изучалась зависимость величины времени реакции (ВР) от индивидуальных особенностей испытуемых (возраст, пол и др.) в задаче трехальтернативного выбора. В первую группу вошли испытуемые в возрасте 20 лет, во вторую — 30 лет, в третью — 40 лет. В табл. 1 (см. с.17) представлены полученные данные. В качестве *зависимой переменной, или результативного признака*, оценивалось ВР на целевой стимул. Предполагалось, что оно будет зависеть от возраста испытуемых, выступающего как причина или *фактор*, вызывающий изменение данного признака. Естественно, что факторов, влияющих на один или несколько признаков, может быть достаточно много. Однако, в эксперименте исследователя интересуют лишь некоторые из них. Они называются *контролируемыми, или изучаемыми факторами*, в отличие от других воздействий, которые тоже оказывают влияние на наблюдаемый признак, но не контролируются. Контролируемость фактора заключается в том, что исследователь задает (или выбирает) определенные градации его изменения, или *уровни фактора*. В нашем примере заданы 3 уровня изменения фактора *возраст*: 20, 30

и 40 лет. Вариант ДА с одним контролируемым фактором называется *однофакторным ДА.* Когда контролируемых факторов два или более, используют *многофакторный ДА.* В нашем исследовании дополнительными изучаемыми факторами могли быть пол, профессиональный опыт, время суток и др. В случае наличия более, чем одного фактора, появляется возможность оценить влияние на зависимую переменную *межфакторного взаимодействия*. Таким образом, ДА как метод обработки эмпирических данных позволяет установить факт влияния фактора на признак или так называемый *главный эффект,* а в том случае, если контролируемых факторов несколько, обнаружить и оценить их *взаимодействие.*

1.2. Линейная модель дисперсионного анализа

В основе ДА лежит обычная линейная модель. Следуя нашему примеру, положим, что $\mu_{общ.}$ — среднее ВР в популяции (генеральной совокупности) взрослых людей, τ_j — обозначает дополнительный компонент, соответствующий изменению возраста испытуемых (т.е. насколько среднее j-го возраста отличается от среднего генеральной совокупности: $\tau_j = \mu_j - \mu_{общ.}$), ε_{ij} — вклад уникальности конкретного (i-го) испытуемого или, как еще говорят, «ошибка» линейной модели. Тогда линейная модель, описывающая результативность четвертого испытуемого первой группы, будет выглядеть следующим образом:

$$x_{4,1} = \mu_{общ.} + (\mu_1 - \mu_{общ.}) + \varepsilon_{4,1} = \mu_{общ.} + \tau_1 + \varepsilon_{4,1} \qquad (1)$$

Это и есть та структурная модель, которая лежит в основе ДА. В более сложных вариантах ДА она может быть расширена, но основная идея остается той же.

Например, допустим, что в среднем по популяции ВР=740 *мс*. Далее, у большой группы двадцатилетних людей, выполнявших подобную задачу, среднее ВР на 30 *мс* больше. И наконец, испытуемый № 4 входит в группу людей, где средний показатель меньше среднего генеральной совокупности на 40 *мс*. Тогда его результативность по данному тесту будет равна:

$x_{4,1} = 740 + 30 - 40 = 730$ *мс.*

Естественно, что мы так подобрали вымышленные статистические оценки, чтобы они соответствовали данным из табл. 1.

1.3. Нулевая гипотеза в ДА

В нашем примере проверяемой нулевой гипотезой будет предположение, что средние величины ВР в различных экспериментальных условиях будут одинаковы. Другими словами, если $\mu_{общ.}$ представляет собой среднее по генеральной совокупности всех испытуемых, которые потенциально могли бы быть участниками нашего эксперимента, а μ_1, μ_2, μ_3 — это популяционные средние для трех экспериментальных условий, то нулевая гипотеза выражается следующим образом:

$$H_0: \mu_{общ.} = \mu_1 = \mu_2 = \mu_3. \qquad (2)$$

Альтернативной гипотезой (H_1) будет предположение о том, что по крайней мере одно среднее отличается от других. Таким образом, в узком смысле ДА — это техника оценки наличия или отсутствия различий между выборочными средними. При оценке ложности H_0 совершенно неважно, что послужило причиной: отличие двух или трех пар средних друг от друга.

Используя при описании нулевой гипотезы термин «генеральная совокупность», или «популяция», мы должны подчеркнуть, что этот термин в контексте ДА означает не множество испытуемых или объектов, а множество *численных значений* определенного признака у этих испытуемых или объектов. Поэтому в ДА мы устанавливаем различие не между популяциями, а между средними значениями исследуемого признака, наблюдаемого (измеряемого) в разных популяциях. В этом смысле корректнее говорить о том, что среднее, полученное по оценкам некоторого признака при одном экспериментальном условии, будет больше или меньше, чем при другом условии. Это уточнение может показаться тривиальным, но это не так. Например, если мы исследуем возрастные особенности цветовых ощущений, то генеральные совокупности людей разного возраста будут, конечно же, отличаться друг от друга множеством характеристик, одна-

ко, совсем не обязательно, что шкальные значения цветовых ощущений будут у них разные.

1.4. Математические допущения ДА

Как метод математической статистики, ДА основывается на ряде допущений о свойствах и параметрах распределения наблюдаемых случайных величин.

Первое допущение ДА требует, чтобы значения признаков, соответствующих каждому уровню контролируемого фактора, были *нормально распределены* вокруг своего среднего. Графически это допущение можно представить в виде нескольких кривых плотностей вероятности, соответствующих нормальному распределению (см. рис. 1). Это, в свою очередь, означает, что каждое из этих распределений будет характеризоваться только двумя параметрами — средним (μ) и дисперсией (σ^2) полученных значений.

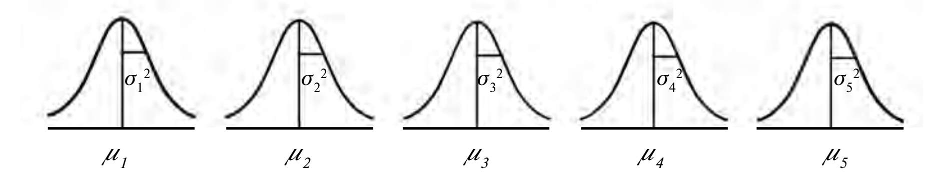


Рис. 1. Графическое представление распределений значений измеряемого признака, полученных при каждом из пяти уровней фактора (см. *Хауэлл*, 1998, с.302)

Второе допущение предполагает равенство дисперсий выборочных распределений, соответствующих каждому уровню контролируемого фактора, т.е. $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$. Это допущение также называют требованием *однородности, или гомогенности,* дисперсий.

Третье допущение касается *независимости* полученных наблюдений. Это предположение означает, что для любых двух наблюдений мы не можем предсказать по значению одного наблюдения значения другого. В нашем эксперименте данное предположение *не выполнялось* бы в том случае, если испытуемые одной группы, записав в ходе опыта последовательность предъявлявшихся стимулов, сообщили бы об этом испытуемым другой группы. Или одна и та же груп-

па испытуемых участвовала бы в эксперименте дважды: например, утром и вечером (фактор: «время суток»). В этом случае более «медленные» испытуемые, участвовавшие в опыте утром, в целом покажут и вечером бо́льшие значения времени реакции, а «быстрые» будут работать вечером также быстрее. Соблюдение независимости наблюдений является важнейшей причиной, обуславливающей в факторном эксперименте необходимость отбора испытуемых в группы случайным образом. Несоблюдение независимости наблюдений может иметь серьезные последствия для результатов ДА.

1.5. Последствия нарушения допущений ДА

Поскольку ДА является точной процедурой статистических расчетов, то, строго говоря, нарушения не могут не повлиять на точность конечных оценок достоверности различия средних. Тем не менее, принимая во внимание, что в реальных данных первые два допущения не всегда строго соблюдаются, математики проводили специальные исследования о влиянии возможных нарушений на результаты ДА. Здесь мы коротко опишем полученные результаты изучения *устойчивости (или робастности)* ДА и дадим соответствующие рекомендации (более подробно об этом см. *Шеффе*, 1980; *Гласс*, *Стэнли*, 1976; *Хауэлл*, 1998).

По данным Г.Шеффе (1962) и др. исследователей (цит. по *Гласс*, *Стэнли*, 1976), влияние неоднородности дисперсий в однофакторном ДА на величину ошибки I рода показало, что вероятность ошибки зависит от размера выборок, числа уровней фактора и реального отношения выборочных дисперсий. Например, если при 3-х уровнях фактора число наблюдений в каждой группе, соответственно, было 9, 5 и 3, а соответствующие дисперсии — 10, 10 и 30 (т.е. находились в отношении 1:1:3), то вероятность ошибки I рода равна 0.17, тогда как экспериментатор традиционно исходил из того, что она равна 0.05. И, следовательно, возрастает вероятность отбросить нулевую гипотезу, когда она действительно верна. Специальные расчеты, однако, показали, что влияние нарушения допущения о равенстве дисперсий может быть компенсировано, когда:

1) объемы выборок равны или отличаются незначительно;
2) используются выборки большого объема.

Таким образом, ДА может быть устойчивым к неоднородности дисперсий и, если дисперсии выборочных групп отличаются не очень значительно и используются выборки равного объема, то его применение оказывается вполне корректным.

Многочисленные исследования показали очень слабую чувствительность ДА к нарушению предположения о нормальности. Следовательно, нарушение предположения о нормальности имеет для ДА небольшое значение, и фактическая вероятность ошибки I рода практически не отличается от устанавливаемой экспериментатором.

1.6. Критерии проверки однородности дисперсий

В том случае, когда есть основания думать о неоднородности групповых дисперсий, следует воспользоваться одним из методов проверки. Поскольку такая проверка носит чисто технический характер, то мы лишь упомянем некоторые из критериев, используемых в современных статистических компьютерных программах, чтобы читатель, зная о них, мог, в принципе, ими воспользоваться.

Критерий Бартлетта сильно чувствителен к предположению о нормальности распределений, поэтому его использование в известном смысле ограничено. Однако, как справедливо утверждают Дж.Гласс и Дж.Стэнли (1976), это позволяет его использовать как средство оценки нормальности. *Критерий Шеффе* не столь чувствителен к нарушению нормальности распределений, поэтому часто используется на практике. Достаточно часто в различных компьютерных программах применяется также *критерий Ливиня*, который не зависит от предположения о нормальности распределений.

1.7. Общая логика ДА

Логические основания, лежащие в основе ДА, просты, и их понимание не требует от психолога серьезной математической подготовки.

Пусть выполняются все допущения ДА и мы имеем несколько выборок с одинаковым числом наблюдений (n_i), соответствующих уровням одного фактора[1]. Пока мы не будем делать никаких предположений о ложности или истинности нулевой гипотезы. Для любого уровня фактора дисперсия значений наблюдаемого признака будет оценкой дисперсии генеральной совокупности, из которой взяты эти значения. Но поскольку мы допустили, что все генеральные совокупности (соответствующие уровням исследуемого фактора) имеют одинаковые дисперсии, то они также являются оценками дисперсии *общей* генеральной совокупности всех значений признака в целом (σ^2_{total}). Переходя к выборочным оценкам дисперсий генеральных совокупностей, заменим символ σ_i^2 на s_i^2:

$\sigma_1^2 = s_1^2,\ \sigma_2^2 = s_2^2,\ \sigma_3^2 = s_3^2,\ \ldots,\ \sigma_k^2 = s_k^2,$ (3)

где знак «=» означает, что σ_i^2 ***оценивается по*** s_i^2, а k — число уровней фактора. Для увеличения надежности оценки дисперсии общей генеральной совокупности усредним дисперсии отдельных выборок, и при принятом нами условии, что $n_1 = n_2 = n_3 = \ldots = n_k$, получим:

$$\sigma^2_{total\,1} = \bar{s}^2_{total} = \frac{s_1^2 + s_2^2 + s_3^3 + \ldots + s_k^2}{k} = s^2_{WG}. \qquad (4)$$

Таким образом, среднее по всем выборочным оценкам дисперсий и будет наилучшей оценкой σ^2_{total1}. В ДА эта *одна из возможных* оценок *общей* дисперсии признака получила название ***внутригрупповой дисперсии*** и обозначается как S^2_{WG} (аббревиатура английского выражения within group — внутригрупповой). Внутригрупповая дисперсия отражает ту часть вариации наблюдаемого признака, которая обусловлена влиянием *случайных* факторов и не зависит от влияния контролируемого фактора.

Подчеркнем, что эта оценка *не зависит* от того, ложна или истинна H_0, т.к. S^2_{WG} вычисляется как простое среднее арифметическое однородных $s^2_1 \ldots s^2_k$.

[1] Как будет показано ниже, для ДА требование равенства числа наблюдений отнюдь не является обязательным.

Теперь обратимся к случаю, когда H_0 верна. Если это так, то k выборок по n наблюдений можно рассматривать как k независимых выборок из *одной и той же* генеральной совокупности наблюдений. В таком случае мы можем использовать *другую* оценку σ^2_{total}. Мы можем оценить общую дисперсию через вариативность групповых средних относительно общего среднего. Из центральной предельной теоремы следует, что вариативность средних, взятых из *одной* генеральной совокупности, равна дисперсии генеральной совокупности, деленной на размер выборки:

$$S^2_{\bar{X}} = \frac{\sigma^2_{total}}{n}. \tag{5}$$

Отсюда следует, что вторая оценка дисперсии — $\sigma^2_{total\,2}$ — может быть вычислена следующим образом:

$$\sigma^2_{total\,2} = S^2_{\bar{X}} \cdot n = S^2_{BG}. \tag{6}$$

Эту (вторую) оценку $\sigma^2_{total\,2}$ называют ***межгрупповой дисперсией*** и обозначают как $\boldsymbol{S^2_{BG}}$ (аббревиатура английского выражения between groups — межгрупповой). В отличие от внутригрупповой, межгрупповая дисперсия отражает не случайную, а систематическую вариацию, т.е. те различия в величине наблюдаемого признака, которые возникают под влиянием изучаемого фактора, уровням которого и соответствуют выделяемые группы.

В свою очередь $S^2_{\bar{X}}$ оценивается следующим образом:

$$S^2_{\bar{X}} = \frac{\sum_k (\bar{X}_{total} - \bar{X}_k)^2}{k-1}, \tag{7}$$

где $\bar{X}_{total}$ — среднее арифметическое всех наблюдений, $\bar{X}_k$ — отдельное выборочное среднее, а k — число выборок.

Таким образом, общая дисперсия оценена двояко: (1) как среднее нескольких выборочных (по числу уровней фактора) дисперсий, и (2) как среднеквадратичное отклонение выборочных средних от общего среднего. Причем очевидно, что первая оценка (S^2_{WG}) не зависит от истинности или ложности

нулевой гипотезы, тогда как вторая (S^2_{BG}) может быть адекватной оценкой общей дисперсии только при условии истинности нулевой гипотезы.

Пример. Рассмотрим два простых числовых примера с вымышленными данными, подобранными специально для иллюстрации того, что происходит, когда нулевая гипотеза истинна или ложна. Допустим, что мы провели эксперимент с тремя группами испытуемых разного возраста (средний возраст в группах — 20, 30 и 40 лет, соответственно). В каждой группе было по 10 человек, выполнявших тест на определение времени реакции (ВР) в условиях трехальтернативного выбора (см. табл. 1, 2).

Случай 1: H_0 верна. В случае истинности H_0 (данные из табл. 1) все средние по трем группам равны: $\bar{X}_1 = \bar{X}_2 = \bar{X}_3$, и каждое значение ВР, взятое из любой выборки, соответствует одной генеральной совокупности.

Оценим *внутригрупповую дисперсию:*

$$S^2_{WG} = \frac{4889 + 6006 + 4694}{3} = 5196 .$$

Далее вычислим *межгрупповую дисперсию*:
сначала

$$S^2_{\bar{X}} = \frac{(770-750)^2 + (725-750)^2 + (755-750)^2}{3-1} = 525 ,$$

и наконец $S^2_{BG} = S^2_{\bar{X}} \cdot n = 525 \cdot 10 = 5250$.

Таким образом, две различные оценки общей дисперсии практически не различаются (сравните: 5196 против 5250), что и соответствует нашему исходному предположению об истинности H_0.

Случай 2: H_0 ложна. Теперь предположим, что во втором эксперименте (данные из табл. 2) в одну из групп (например, в группу 2) попали испытуемые, ранее уже участвовавшие в таком же опыте по измерению ВР. Таким образом, это были хорошо тренированные испытуемые, поэтому мы можем ожидать, что у них в целом значения ВР будут меньше, чем в двух других группах. Для симуляции подобного эффекта мы уменьшили все значения ВР испытуемых группы 2 на 50 *мс*.

Таблица 1

Данные эксперимента по определению ВР для случая: H_0 истинна

	ВР группы 1, *мс*	ВР группы 2, *мс*	ВР группы 3, *мс*	
	630	690	760	
	780	810	760	
	830	740	860	
	730	790	810	
	680	690	660	
	780	740	810	
	830	690	660	
	830	840	810	
	780	690	710	
	830	570	710	
n	10	10	10	N_{total} = 30
$\overline{X}_i$	770	725	755	$\overline{X}_{total}$ = 750
s_1^2	4889	6006	4694	s_{total}^2 = 5200

Таблица 2

Данные эксперимента по определению ВР для случая: H_0 ложна

	ВР группы 1, *мс*	ВР группы 2, *мс*	ВР группы 3, *мс*	
	630	640	760	
	780	760	760	
	830	690	860	
	730	740	810	
	680	640	660	
	780	690	810	
	830	640	660	
	830	790	810	
	780	640	710	
	830	520	710	
n	10	10	10	N_{total} = 30
$\overline{X}_i$	770	675	755	$\overline{X}_{total}$ = 733 *мс*
s_i^2	4889	6006	4694	s_{total}^2 = 6637

Так же оценим и в этом примере *внутригрупповую дисперсию*:

$$s_{WG}^2 = \frac{4889 + 6006 + 4694}{3} = 5196 .$$

Очевидно, что внутригрупповая дисперсия не должна была измениться.

Вычислим *межгрупповую дисперсию*:
сначала

$$s_{\bar{X}}^2 = \frac{(770-733)^2 + (675-733)^2 + (755-733)^2}{3-1} = 2608.5,$$

и далее $s_{BG}^2 = s_{\bar{X}}^2 \cdot n = 2608.5 \cdot 10 = 26085$.

По сравнению с предыдущим примером *межгрупповая* дисперсия (в отличие от внутригрупповой!) возросла значительно, что отразило влияние различий между тремя групповыми средними.

Таким образом, данные примеры показывают, и это мы подчеркиваем особо, что s_{BG}^2 является не просто еще одной оценкой дисперсии генеральной совокупности (σ_{total}^2), но кроме этого и оценкой разброса самих выборочных средних. Другими словами, средние трех экспериментальных групп отличаются прежде всего не в силу случайной вариации величины ВР, а по причине более успешного выполнения данного теста тренированными испытуемыми группы 2.

В заключение еще раз дадим краткую характеристику общей логики ДА. Для проверки H_0 мы должны вычислить две оценки дисперсии генеральной совокупности: s_{WG}^2, которая не зависит от истинности или ложности нулевой гипотезы, и s_{BG}^2, зависящую от истинности нулевой гипотезы. Когда обе эти оценки хорошо согласуются, у нас нет оснований отвергать H_0. В случае их сильного расхождения (когда s_{BG}^2 намного больше s_{WG}^2) одно или несколько средних отличаются друг от друга, и мы можем предположить, что варьируемый фактор внес существенный вклад в их различия. Тогда H_0 должна быть отвергнута.

Глава 2

Однофакторный дисперсионный анализ

В двух предыдущих примерах (см. гл. 1) мы имели дело с анализом данных типичного однофакторного эксперимента. Обратимся к этим данным еще раз и проделаем ДА до конца, т.е. строго проверим нулевую гипотезу и оценим ее статистическую достоверность. Для этого рассмотрим понятия *F*-отношения и степени свободы.

2.1. Процедура оценки *F*-отношения

F-отношение, или критерий Фишера — это статистика, рассчитываемая в ДА. Этот показатель характеризует сравнение дисперсии, обусловленной вариацией самих групповых средних относительно общего среднего, с дисперсией, обусловленной вариацией признака внутри каждой отдельной группы относительно среднего по группе. Однако, при вычислении *F*-отношения прямо не используются оценки S^2_{BG} и S^2_{WG}, введенные нами ранее, а применяются несколько другие оценки, основанные на расчете соответствующих *сумм квадратов* и соотнесении их с соответствующим каждому источнику вариации числом *степеней свободы*. Рассмотрим все необходимые вычисления подробно.

Напомним, что выборочная дисперсия признака вычисляется как средняя сумма квадратов отклонений каждого значения выборки от среднего:

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{N-1} = \frac{\sum d_i^2}{N-1}, \qquad (8)$$

отсюда $\sum d_i^2 = (N-1) \cdot S^2$, (9)

где $\sum d_i^2$ и обозначает *сумму квадратов* соответствующих отклонений, и, естественно, также характеризует величину вариации признака относительно среднего. Отметим, что для получения несмещенной оценки выборочной дисперсии вместо N в формуле используют величину $(N - 1)$.

Обратимся вновь к данным примера 1 (см. табл. 3).

Таблица 3

Результаты вычислений по данным примера 1

	Группа 1	Группа 2	Группа 3	
n	10	10	10	$N_{total} = 30$
$\overline{X}_i$	770	725	755	$\overline{X}_{total} = 750$
s_i^2	4889	6006	4694	$S_{total}^2 = 5200$
$\sum d_{total}^2 = 150800$, $df_{total} = 29$				$\hat{S}_{BG}^2 = 5250$
$\sum d_{BG}^2 = 10500$, $df_{BG} = 2$				$\hat{S}_{WG}^2 = 5389$
$\sum d_{WG}^2 = 140300$, $df_{WG} = 27$				$F = 1.01$

Используя последнюю формулу, легко рассчитать *общую* по всем трем группам сумму квадратов:

$$\sum d_{total}^2 = (N_{total} - 1) \cdot S_{total}^2 = 29 \cdot 5200 = 150800.$$

Межгрупповую (или факторную) сумму квадратов (в отличие от общей суммы квадратов) рассчитывают по следующей формуле:

$$\sum d_{BG}^2 = n \cdot \sum_{j=1}^{N} \overline{X}_j^2 - \frac{\left(\sum_{j=1}^{N} X_i\right)^2}{N}, \qquad (10)$$

где n — число наблюдений в каждой группе, j — знак суммирования квадратов средних по k группам, а i — суммирования значений всех наблюдений в отдельности. Подставляя численные значения, получим:

$$\sum d_{BG}^2 = 10 \cdot (770^2 + 725^2 + 755^2) - \frac{(22500)^2}{30} =$$

$$= 16885500 - 16875000 = 10500.$$

Внутригрупповую (или остаточную) сумму квадратов обычно вычисляют как разность между общей и межгрупповой суммой квадратов[1]:

$$\sum d_{WG}^2 = \sum d_{total}^2 - \sum d_{BG}^2 = 150800 – 10500 = 140300.$$

Далее следует определить число *степеней свободы* для межгрупповой и внутригрупповой вариаций наблюдений. Математически строгое введение понятия степени свободы достаточно трудно, поэтому мы дадим лишь его упрощенную трактовку, в соответствии с которой для некоторого множества наблюдений число степеней свободы соответствует числу не зависящих друг от друга наблюдений. Например, нам известно, что сумма пяти чисел равна 15. Первые четыре числа могут быть любыми: 1, 2, 3 и 4 или 10, 1, 20 и 3. Но, если сумма первых четырех чисел равна 10, а вторых — 34, то, следовательно, в первом случае последнее пятое число *должно быть* только «5», а во втором равняться «-19».

Таким образом, последнее (пятое) число уже не является независимым. Следовательно, в данном числовом примере имеется 4 степени свободы, т.е. на единицу меньше общего количества чисел. Вернувшись к нашему примеру 1, где в исследовании участвовало 30 испытуемых, можно определить число степеней свободы для оценки *общей* дисперсии. Оно будет равно:

$$df_{total} = N — 1 = 30 — 1 = 29. \qquad (11)$$

Обычно число степеней свободы обозначают *df* (аббревиатура от английского выражения degrees of freedom).

Для оценки *межгрупповой* вариации число степеней свободы оценивается так:

$$df_{BG} = k — 1 = 3 — 1 = 2, \qquad (12)$$

т.е. равно числу уровней фактора (в нашем примере — числу групп) минус 1.

---

[1] Кроме того, $\sum d_{WG}^2$ можно вычислить как общую сумму квадратов, состоящую из трех сумм квадратов соответствующих групп наблюдений:

$$\sum d_{WG}^2 = \sum d_{WG_1}^2 + \sum d_{WG2}^2 + \sum d_{WG_3}^2 = n \cdot s_1^2 + n \cdot s_2^2 + n \cdot s_3^2.$$

Для оценки *внутригрупповой* вариации число степеней свободы рассчитывается следующим образом:

$$df_{WG} = N - k = 30 - 3 = 27. \qquad (13)$$

Число 27 можно проинтерпретировать следующим образом: каждая группа, состоящая из 10 наблюдений, «теряет» по 1 степени свободы, следовательно, данные по каждой группе имеют 9 степеней свободы, а групп — 3, поэтому $9 \times 3 = 27$. Для контроля правильности вычислений следует помнить, что сумма степеней свободы для внутри- и межгрупповой вариации всегда равняется общему числу степеней свободы, т.е.

$$df_{total} = df_{BG} + df_{WG} = 2 + 27 = 29. \qquad (14)$$

После вычисления сумм квадратов и расчета числа степеней свободы можно найти оценки межгрупповой ($\hat{S}^2_{BG}$) и внутригрупповой ($\hat{S}^2_{WG}$) дисперсий, входящих в F-отношение. Для этого полученные суммы квадратов делятся на соответствующее им число степеней свободы:

$$\hat{S}^2_{BG} = \frac{\sum d^2_{BG}}{k-1}, \quad \hat{S}^2_{WG} = \frac{\sum d^2_{WG}}{N-k}. \qquad (15)$$

Искомая статистика дисперсионного анализа — F-критерий — выглядит как простое отношение этих двух дисперсий:

$$F = \frac{\hat{S}^2_{BG}}{\hat{S}^2_{WG}}. \qquad (16)$$

Вычислим величину F-отношения для нашего примера 1:

$$F = \frac{5250}{5196} = 1.01.$$

Еще раз напомним, что величина F представляет собой отношение оценки общей дисперсии, обусловленной воздействием некоторого общего для всей совокупности наблюдений фактора, к дисперсии, обусловленной вариацией наблюдений внутри отдельных групп, соответствующих уровням этого фактора.

Далее необходимо сравнить найденное значение F-отношения с табличным значением (см. табл. Приложения). Иначе говоря, с помощью таблиц F-распределения определяется вероятность получения данной величины F путем его сопоставления с соответствующим выборочным распределением F-значений. При пользовании таблицами исследователь выбирает определенную доверительную вероятность (т.е. вероятность ошибки I рода), а затем ищет табличное значение F-критерия с учетом определенного ранее числа степеней свободы. Выберем в табл. Приложения уровень значимости $p = 0.05$, и на пересечении столбца 2 (т.е. две степени свободы у числителя) и строки 30 (это самое близкое значение к числу степеней свободы знаменателя) считаем критическое значение F-критерия — 3.32. Найденное нами значение F меньше критического, следовательно у нас нет оснований отвергать нулевую гипотезу. Это означает, что исследуемый фактор (в нашем примере это возраст испытуемых) не оказывает статистически достоверного влияния на величину времени реакции.

Результаты дисперсионного анализа обычно записываются в таблицу *стандартного* вида (см. табл. 4). В последнем столбце табл. 4 указывается уровень значимости, при котором нулевая гипотеза может быть отвергнута.

Таблица 4

Результаты дисперсионного анализа данных из примера 1

Источник вариации	Степени свободы	$\sum d^2$	$\hat{S}^2$	F	P
Межгрупповая	27	10500	5250	1.01	>0.05
Внутригрупповая	2	140300	5196		
Общая	29	150800			

Далее вычислим величину F-отношения для примера 2 (см. табл. 5). Полученное в примере 2 значение F больше критического ($F_{кр.} = 3.32$), следовательно нулевая гипотеза может быть отвергнута на уровне значимости $p < 0.05$.

Таблица 5

Результаты вычислений по данным из примера 2

	Группа 1	Группа 2	Группа 3	
n	10	10	10	N_{total} = 30
$\overline{X}_i$	770	675	755	$\overline{X}_{total}$ =733 *мс*
S_i^2	4889	6006	4694	S_{total}^2 = 6637
$\sum d_{total}^2$ =192467 , df_{total} = 29 $\sum d_{BG}^2$ =52167 , df_{BG} = 2 $\sum d_{WG}^2$ =140300 , df_{WG} = 27			$\hat{S}_{BG}^2$ = 26083 $\hat{S}_{WG}^2$ = 5196 F = 5.02	

Строго говоря, это означает, что по крайней мере одно из средних отличается от других. В нашем примере это отличие обусловлено влиянием фактора «небрежности экспериментатора», случайно пригласившего дважды принять участие в эксперименте одних и тех же испытуемых. Для того, чтобы конкретно понять, какое (или какие) же среднее отличается от других, используют процедуры попарного сравнения средних.

2.2. Процедуры множественного сравнения средних в ДА

Казалось бы, что большой проблемы с проведением такого сравнения нет — достаточно воспользоваться *t-критерием* Стьюдента. Однако, оказывается, что с ростом количества сравниваемых средних происходит значительное увеличение вероятности ошибки I рода. Допустим, что мы сравниваем между собой 5 средних, и в силу *случайной ошибки* в одной из выборок, например, во 2-ой, ее среднее оказалось выше, чем в других. Для того, чтобы определить значимость различий между средними, нам нужно произвести попарные сравнения всех средних: $\overline{X}_1 - \overline{X}_2$, $\overline{X}_1 - \overline{X}_3$, $\overline{X}_1 - \overline{X}_4$ и т.д. — всего $\frac{n(n-1)}{2}$ или 10 пар сравнений. Из них в 5 случаях мы обнаружим, что $\overline{X}_2$ больше остальных. Следовательно, в 50% сравнений мы допустим ошибку I рода, т.е. отвергнем нулевую

гипотезу, когда она в действительности верна, и сделаем заключение, что $\overline{X}_2$ отличается от среднего по совокупности в целом. Таким образом, при увеличении количества сравниваемых средних *в силу случайности выборки* в принципе возрастает вероятность совершить ошибку I рода. Поэтому однородность всей совокупности, оцениваемой по нескольким выборочным средним, оказывается *сильно зависимой* от средней всего лишь *одной* выборки. В ДА для решения этой задачи используется другой подход, когда в оценку значимости различий между средними вносят вклад *сразу все* выборочные средние. Коротко данный подход можно выразить так: не слишком ли велика *дисперсия средних*, чтобы предположить, что все они являются средними выборок, взятых из одной генеральной совокупности?

Одна из таких процедур была предложена самим Р.Фишером и получила название множественного *t-критерия* или теста минимально значимых различий Фишера[1]. Этот метод оценки парных различий между средними наряду с методами Шеффе и Бонферрони известен как наиболее либеральный из всех методов. Использование *t-коэффициента Фишера* оправдано лишь в том случае, если в результате проведенного ДА оценка *F*-отношения оказалась значима, а также при однородности дисперсий сравниваемых выборок. Расчет множественного *t*-критерия похож на вычисление *t*-критерия Стьюдента, однако вместо оценки дисперсии каждой выборки в отдельности в формуле участвует оценка *внутригрупповой* дисперсии, зависящая от вариабельности данных внутри ***всех*** групп, а не только двух сравниваемых:

$$t_F = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\dfrac{\hat{S}^2_{WG}}{n_i} + \dfrac{\hat{S}^2_{WG}}{n_j}}}, \tag{17}$$

где n_i и n_j — размер сравниваемых выборок (в нашем примере они равны).

[1] Иногда данный коэффициент называют также защищенным *t*-критерием.

Используя данные из примера 2, оценим достоверность различий средних ВР по первой и второй группам:

$$t_{F_{1-2}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\hat{s}^2_{WG}}{n_1} + \frac{\hat{s}^2_{WG}}{n_2}}} = \frac{95}{32.2} = 2.95.$$

Для оценки достоверности полученных различий используем таблицу критических значений двухстороннего t-критерия Стьюдента и найдем критическое значение t-критерия при уровне достоверности 0.05 и количестве степеней свободы 18: $t_{кр.} = 2.1$. Таким образом, мы (как и ожидалось) обнаружили статистически достоверные различия между сравниваемыми групповыми средними. Аналогичная оценка различий между средними первой и третьей групп (см. табл. 6) не обнаружила достоверных различий.

Таблица 6

Результаты парных сравнений средних из примера 2

Сравнение средних	По Фишеру	По Шеффе	По Бонферрони
Группа 1 — группа 2: $t = 2.95$	$t_{F,\ кр.} = 2.1$	$t_{Sh,\ кр.} = 2.58$	$t_{B,\ кр.} = 2.43$
Группа 1 — группа 3: $t = 0.47$	$t_{F,\ кр.} = 2.1$	$t_{Sh,\ кр.} = 2.58$	$t_{B,\ кр.} = 2.43$
Статистический вывод:	1) ***есть*** различие между средними 1 и 2; 2) ***нет*** различия между средними 1 и 3.	1) **есть** различие между средними 1 и 2; 2) ***нет*** различия между средними 1 и 3.	1) **есть** различие между средними 1 и 2; 2) ***нет*** различия между средними 1 и 3.

Для более строгой оценки различий между средними используют также *тест Шеффе*. При использовании метода Шеффе требование статистической достоверности F-критерия не является обязательным, что делает его более универсальным. При оценке различий по Шеффе (t_{Sh}) используется та же самая формула множественного t-критерия:

$$t = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\frac{\hat{s}^2_{WG}}{n_i} + \frac{\hat{s}^2_{WG}}{n_j}}}. \qquad (18)$$

Однако, достоверность различия между средними оценивается иначе. Полученное значение (t_{Sh}) сравнивается с критическим, которое, в свою очередь, находится на основе критических значений F-критерия:

$$t_{Sh,\,кр.} = \sqrt{F-(k-1)}, \qquad (19)$$

где k — число уровней фактора.

При сравнении групп 1 и 2 величина t_{Sh} будет также равна 2.95. По таблицам критерия Фишера находим соответствующее $F_{кр.} = 3.32$, следовательно $t_{Sh,\,кр.} = \sqrt{3.32 \cdot (3-1)} = 2.58$. Таким образом, при использовании критерия Шеффе мы тоже обнаружили достоверность различия между средними первой и второй групп, хотя разница между найденной величиной t_{Sh} и соответствующим критическим значением уже не столь велика (0.37 против 0.85).

Критерий Бонферрони (иногда этот критерий называют также тестом Данна — Dunn's test — по имени ученого, формализовавшего его применение) можно считать хорошим компромиссом между двумя рассмотренными выше критериями. Для его использования рассчитывается множественный t-критерий. Однако, критическое значение t-критерия оценивается несколько сложнее. Оно зависит от количества сравниваемых пар средних. Нулевая гипотеза отвергается на уровне достоверности, вычисляемом как p/c, где p — принятый уровень достоверности (например: 0.05 или 0.01), а c — количество проводимых сравнений. Например, если мы хотим сравнить 2 пары средних, то нулевую гипотезу следует отвергать на уровне 0.05/2=0.025, а если сравниваются 3 пары, то на уровне 0.05/3=0.0167 и т.д. Естественно, что для использования критерия Бонферрони применяются не стандартные, а специальные таблицы (см., напр., *Хауэлл*, 1998) или аппроксимирующие формулы.

Расчеты, суммированные в таблице 6, показывают, что тест Фишера оказывается самым либеральным, тест Шеффе — самым строгим, а тест Бонферрони занимает промежуточное положение. Таким образом, при принятии статистического решения на грани достоверности оцениваемых сравнений следует учитывать подобные возможности.

Некоторые авторы (см. *Хауэлл*, 1995) предлагают еще один способ расчета критерия Шеффе: t-коэффициент Стьюдента возводится в квадрат и сравнивается с удвоенной величиной критического значения F-критерия.

Следует отметить, что существуют и в различных статистических пакетах реально используются и другие критерии множественного сравнения средних. Например, в системе SPSS 8.0 их 18. Для более направленного и осмысленного выбора следует обратиться к специальной литературе (см., напр., *Хауэлл*, 1998). Кроме того, необходимо знать, что перед сравнением средних соответствующие выборки стоит оценить на однородность дисперсий, и, в случае обнаружения неоднородности, воспользоваться теми критериями, которые не требуют выполнения этого условия. К таковым относятся, например, критерии Тамхана, Даннета и Геймса—Хауэлла.

2.3. ДА с выборками неравного размера

В рассмотренных выше примерах число испытуемых в каждой группе, т.е. число наблюдений (измерений) в каждой выборке, было одинаковым. Это было сделано лишь из соображений удобства и простоты вычислений. В реальной практике обработки экспериментальных данных вследствие разных причин[1] часто случается, что ДА подвергаются данные, в которых выборки, соответствующие различным уровням одного фактора, оказываются неодинакового размера. В этом случае изменения в вычислениях касаются лишь одной формулы — формулы для оценки межгрупповой суммы квадратов. Данная модификация касается лишь одного

[1] Нередко испытуемые выбывают из эксперимента или результаты нескольких опытов оказываются негодными.

параметра — числа наблюдений в отдельных выборках, поэтому множитель n следует заменить на n_j и внести его под знак суммы:

$$\sum d_{BG}^2 = \sum_{j=1}^{k} \overline{X}_j^2 \cdot n_j - \frac{\left(\sum_{i=1}^{N} X_i\right)^2}{N}. \tag{20}$$

2.4. Оценка силы факторного эффекта

В том случае, если установлено статистически достоверное влияние контролируемого фактора на зависимую переменную, возможно количественно оценить силу этого влияния. Как правило, силу факторного влияния измеряют с помощью различных показателей, оценивающих отношение межгрупповой дисперсии (т.е. дисперсии, обусловленной влиянием фактора) к общей дисперсии. Существует несколько подобных показателей.

Один из простейших — η^2 (*эта квадрат*):

$$\eta^2 = \frac{\sum d_{BG}^2}{\sum d_{total}^2}. \tag{21}$$

Очевидно, что при полном отсутствии факторного эффекта η^2 будет равно 0, а при значительном влиянии фактора — стремиться к 1.

В примере 1, где не было обнаружено основного эффекта, этот показатель будет равен: $\eta_{1пр.}^2$ = 10500/150800 = 0.07. Ясно, что эта величина незначительно отличается от 0 — всего 7% от общей дисперсии. В примере 2, где было достоверно обнаружено наличие факторного эффекта, $\eta_{2пр.}^2$ = 52167/192467 = 0.27. Эта величина почти 4 раза больше, чем в первом примере — 27% вариабельности данных обусловлено факторным эффектом. Этот показатель часто используется в качестве меры силы влияния фактора, хотя, как показывают специальные расчеты, часто его переоценивает. Более точной, несмещенной оценкой является другая статистика — ω^2 (*омега квадрат*):

$$\omega^2 = \frac{\sum d_{BG}^2 - (k-1)\cdot \hat{s}_{WG}^2}{\sum d_{total}^2 + \hat{s}_{WG}^2}, \tag{22}$$

где k — число уровней фактора.

Вычислим и этот показатель для двух наших примеров:

$$\omega_{1пр.}^2 = \frac{10500-(3-1)\cdot 5196}{150800-5196} = \frac{108}{145604} = 0.0074,$$

$$\omega_{2пр.}^2 = \frac{52167-(3-1)\cdot 5196}{192467-5196} = \frac{41775}{187271} = 0.23.$$

Видно, что величина ω^2 в обеих случаях меньше, чем η^2.

Во многих руководствах по ДА и в некоторых компьютерных программах для оценки силы влияния фактора используется *показатель Снедекора*:

$$h^2 = \frac{\hat{s}_{BG}^2 - \hat{s}_{WG}^2}{\hat{s}_{BG}^2 + (n-1)\hat{s}_{WG}^2}, \tag{23}$$

где n — численность наблюдений в отдельных группах. Если в группах неодинаковое число наблюдений, то n будет вычисляться сложнее:

$$n = \frac{1}{k-1}\left(N - \frac{\sum n_i^2}{N}\right), \tag{24}$$

где k — число уровней фактора, а N — общее число наблюдений.

Вычислим h^2 Снедекора для двух наших примеров:

$$h^2{}_{1пр.} = \frac{5250-5389}{5250+(10-1)\cdot 5389} = \frac{-139}{53751} = -0.0026,$$

$$h^2{}_{2пр.} = \frac{26083-5196}{26083+(10-1)\cdot 5196} = \frac{20887}{72847} = 0.287.$$

Как видно, оценки факторного влияния, рассчитанные тремя различными способами, в целом дают сходные ре-

зультаты: для первого примера они близки к нулю, для второго составляют 23—27%. Последние свидетельствуют о существенном вкладе межгрупповой дисперсии в общую дисперсию.

2.5. О двух моделях дисперсионного анализа

При обработке данных однофакторного эксперимента наиболее разработанными и поэтому распространенными считаются две модели. Их различие обусловлено спецификой *планирования* самого эксперимента. В модели ДА с *фиксированными эффектами* исследователь *намеренно* устанавливает строго определенные уровни изучаемого фактора. Термин «фиксированный эффект» в данном контексте имеет тот смысл, что самим исследователем фиксируется количество уровней фактора и различия между ними; при повторении эксперимента он (или другой исследователь) выберет те же самые уровни фактора. В модели со *случайными эффектами* уровни значения фактора выбираются исследователем *случайно* из широкого диапазона изменений фактора, и при повторных экспериментах, естественно, этот диапазон будет другим.

Таким образом, данные модели отличаются между собой способом выбора уровней фактора, что, очевидно, в первую очередь влияет на возможность обобщения полученных экспериментальных результатов. Для ДА однофакторных экспериментов различие этих двух моделей не столь существенно, однако в многофакторном ДА оно может оказаться весьма важным.

Контрольные вопросы

1. Что понимается в ДА под зависимой переменной и фактором?
2. Что такое уровень контролируемого фактора? Чем отличаются модели ДА с фиксированными и случайными эффектами?
3. Каковы различия между однофакторным и многофакторным ДА?

4. Что такое главный эффект и межфакторное взаимодействие?

5. Какая структурная модель лежит в основе ДА?

6. Какая нулевая (H_0) гипотеза проверяется в ДА? Сформулируйте альтернативную гипотезу (H_1) ДА.

7. Сформулируйте три основных математических допущения ДА.

8. Каковы последствия нарушения допущений ДА?

9. Какие статистические критерии используются для оценки однородности дисперсий? Каким образом можно компенсировать влияние нарушения допущения об однородности дисперсий?

10. Что такое внутригрупповая дисперсия и как она оценивается?

11. Что такое межгрупповая дисперсия и как она оценивается?

12. Что оценивается с помощью *F*-отношения, какая гипотеза проверяется? Как рассчитывается *F*-отношение?

13. Каким образом вычислить межгрупповую и внутригрупповую суммы квадратов?

14. Что такое степень свободы и как определить число степеней свободы при оценке межгрупповой ($\hat{S}^2_{BG}$) и внутригрупповой ($\hat{S}^2_{WG}$) дисперсий?

15. Как определить статистическую значимость полученного *F*-отношения?

16. С какой целью применяют методы множественного сравнения средних? Почему использование *t*-критерия Стьюдента считается не вполне корректным?

17. Какие критерии парных сравнений используются на практике? В чем их особенность по сравнению с *t*-критерием Стьюдента?

18. Как оценивают силу факторного влияния? В чем смысл используемых показателей?

2.6. Работа с однофакторным дисперсионным анализом в статистической системе STADIA 6.0

Обработаем данные примера, взятого из учебного пособия С.А.Шапкина (1997). Автор приводит данные исследования феномена так называемой *«выученной беспомощности»*. Как извест-

но, данный феномен, исследуемый в рамках психологии личности, заключается в снижении у человека способности к решению задач, если перед этим он подвергался переживанию неуспеха в сходных задачах. В эксперименте приняли участие 3 группы испытуемых. Первая группа в предварительной серии проб пыталась решать *неразрешимые анаграммы,* в тестовой серии этой группе предъявлялись *также анаграммы*, но уже вполне разрешимые. Второй группе испытуемых в предварительной серии также предлагались неразрешимые анаграммы, но в тестовой серии были предложены *другие (разрешимые) задачи* — на установление закономерностей. Третья группа испытуемых (контрольная) в предварительной серии не подвергалась переживанию неуспеха — им предлагались разрешимые анаграммы, а в тестовой серии они решали оба типа задач (анаграммы и задачи на установление закономерностей). Таким образом, мы имеем дело с однофакторным экспериментом, где зависимая переменная — время решения экспериментальной задачи, а фактором является условие решения задач в тестовой серии, задаваемое в предварительной серии (3 уровня фактора: «однородное» условие, «неоднородное» условие и контрольное условие). В таблице 7 приведены экспериментальные данные. Первый столбец — порядковый номер испытуемого. Во втором — экспериментальное условие (1, 2 и 3, соответственно). В третьем — среднее время решения задачи в тестовой серии.

Таблица 7

Данные, полученные в эксперименте по изучению феномена «выученной беспомощности»

№ испытуемого	Условие решения	Время решения, *с*
1	1	97
2	1	89
3	1	101
4	1	115
5	2	111
6	2	112
7	2	87
8	2	80
9	3	49
10	3	66
11	3	72
12	3	77

Для обработки данных в системе Stadia они заносятся в электронную таблицу в виде псевдоматрицы, т.е. столбцы не обязательно должны быть одинаковой длины как в нашем примере. При вводе чисел переменные (столбцы электронной таблицы) должны соответствовать различным уровням исследуемого фактора, а строки — отдельным наблюдениям по каждому из уровней фактора. В таблице 8 данные представлены в том виде, в котором они должны быть введены в электронную таблицу.

Таблица 8

Ввод данных в электронную таблицу статистической системы Stadia

	Однород.	Неоднород.	Контрол.
1	97	111	49
2	89	112	66
3	101	87	72
4	115	80	77

Для удобства работы переменным можно присвоить характерные имена, содержащие до 8 букв, например, как это сделано в таблице 8. После входа в меню статистических методов (**F9**) следует выбрать *раздел «Дисперсионный анализ»*, а в нем опцию — *«Однофакторный»*. После этого в окне *«Анализ переменных»* выбрать нужные вам переменные (в нашем случае их 3) или, если лишних нет, нажать на кнопку *«Все»*. После нажатия на кнопку *«Утвердить»* следует выбрать *«параметрический»* метод ДА. Результаты анализа представляют собой стандартный набор рассчитанных показателей (см. табл. 9). Кроме того, STADIA в словесном виде формулирует результат проверки нулевой гипотезы. В конце таблицы результатов выдаются значения параметров однофакторной модели: общее среднее, отклонение от среднего, вызванное влиянием каждого из уровней фактора (эффект1—3), и соответствующие им значения доверительных интервалов. В случае наличия факторного эффекта (в нашем случае он имеет место) выдается таблица парных сравнений средних по Шеффе, где приводятся следующие параметры (по столбцам): разность средних значений, размах довери-

тельного интервала разности, уровень значимости нулевой гипотезы об отсутствии различий между средними и факт принятия альтернативной гипотезы. Далее исследователь может продолжить сравнение средних путем объединения их в отдельные группы или закончить анализ.

Как следует из приведенных результатов ДА (см. табл. 9), мы обнаружили статистически значимое влияние фактора «условие решения задачи» (p = 0.0096) на зависимую переменную «время решения задачи».Как показывает результат парных сравнений групповых средних по Шеффе, факторный эффект обусловлен превышением среднего времени решения задач в тестовой серии при «однородном» и «неоднородном» условиях по сравнению с «контрольным» условием — 34.5 и 31.5, соответственно. Различие среднего времени решения при «однородном» и «неоднородном» условиях оказалось статистически незначимым (3; p = 0.951).

Таблица 9

Пример выдачи результатов дисперсионного анализа статистической системой STADIA 6.1/prof

1-ФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ. Файл: learnhelp.std
параметрический

Источник	Сум.квадр	Ст.своб	Ср.квадр	Сила влияния
Факт.1	2.92E3	2	1.46E3	0.134
Остат.	1.61E3	9	179	
Общая	4.53E3	11	412	

F(фактор1)=8.17, Значимость=0.0096, степ.своб = 2,9

Гипотеза 1: <Есть влияние фактора на отклик>

Параметры модели:
Среднее = 88, доверит.инт.=43.6
Эффект1 = 12.5, доверит.инт.=81.9
Эффект2 = 9.5, доверит.инт.=81.9
Эффект3 = -22, доверит.инт.=81.9

Парные сравнения Шеффе

Переменные	Разность	Интервал	Значим.	Гипотеза H1
1-2	3	27.5	0.951	
1-3	34.5	27.5	0.0167	**Да**
2-3	31.5	27.5	0.0266	**Да**

2.7. Работа с однофакторным дисперсионным анализом в статистической системе SPSS 8.0.1[1]

Данные заносятся в электронную таблицу *Редактора Данных* (SPSS Data Editor) в виде двух столбцов (переменных) одинаковой длины: в первый заносятся значения исследуемой переменной, полученные по каждому из N наблюдений, во второй — значения уровня исследуемого фактора (для k фиксированных уровней — это числа от 1 до k), соответствующие отдельным наблюдениям[2]. Таким образом, количество строк равно сумме *всех* наблюдений или сумме наблюдений, полученных по каждому из k уровней фактора. Если однофакторный ДА выполняется одновременно для нескольких переменных, то электронная таблица, соответственно, может содержать большее количество столбцов.

В меню статистических методов *Статистика* (Statistics) следует выбрать группу методов *Сравнение средних* (Compare Means), а в ней — *Однофакторный дисперсионный анализ* (One-Way ANOVA). В появившемся окошке слева будет представлен список всех переменных (столбцов), включенных в электронную таблицу. Из них нужно выбрать анализируемые переменные и перенести их в окно *Список зависимых* (Dependent List). Затем следует указать, в какой переменной заданы значения уровней исследуемого фактора и перенести ее в окно *Фактор* (Factor). Дополнительно можно заказать расчет дополнительных статистических показателей, нажав на кнопку *Параметры* (Options): к результатам добавятся показатели описательной статистики по каждой переменной и будет произведена проверка дисперсий сравниваемых групп на гомогенность по критерию Ливиня.

---

[1] При описании приемов работы с SPSS мы используем вариант русской версии данной статистической системы — SPSS 8.0.1 RUS. За исключением однофакторного ДА, по сравнению с предыдущими версиями системы в процедурах ДА были сделаны существенные изменения. Для тех, кто пользуется англоязычной версией системы, мы приводим в скобках оригинальные названия пунктов меню.

[2] Форма представления данных в таблице 7 полностью соответствует тому, что требуется для их обработки в системе SPSS.

Кроме того SPSS позволяет оценить так называемые *априорные контрасты*. Если исследователь предполагает, что сравниваемые групповые средние представляют собой некоторый *упорядоченный* (в соответствии со значением уровня исследуемого фактора) ряд, то можно попробовать оценить характер соотношения между этими средними. Таким образом, расчет контрастов используется для выявления взаимосвязей между средними. Фактически контраст — это линейная комбинация нескольких средних ($\bar{X}_i$) с коэффициентами a_i:

$$a_1\bar{X}_1 + a_2\bar{X}_2 + ... + a_k\bar{X}_k = 0, \qquad (25)$$

где $a_1 + a_2 + ... + a_k = 0$.

Таким образом, задав вид полинома, можно оценить характер соотношения между средними. Для этого нужно в опции *Контрасты* (Contrasts) выбрать пункт *Полином* (Polinomial) и в окошке *Степень* (Degree) указать вид проверяемой зависимости — линейной, квадратической, кубической или др. Исследователь может также задать и произвольный набор коэффициентов. SPSS предполагает, что коэффициент, заданный первым, соответствует минимальному уровню фактора, второй — второму по величине и т.д. Сумма коэффициентов должна быть равна 0.

До или сразу после проведения однофакторного ДА исследователь может выбрать критерии для множественного сравнения средних или, как еще говорят, воспользоваться *апостериорными критериями сравнения, или оценить апостериорные контрасты*. Для это следует нажать на кнопку *Апостериорные* (Post Hoc). SPSS представляет 18 методов множественного сравнения средних. Можно выделить 2 типа таких статистических тестов: проверяющие разности между каждой парой средних и выделяющие однородные группы средних, которые не отличаются друг от друга. Наиболее часто используемыми являются методы *Бонферрони и Тьюки*. Кроме того, метод Тьюки также выделяет однородные группы сред-

них. При большом количестве сравниваемых средних более чувствительным считается метод Тьюки, при малом — более предпочтителен метод Бонферрони. Оба метода требуют равенства групповых дисперсий.

SPSS предоставляет 4 метода, допускающих неравенство дисперсий: Геймса—Хауэлла, T2 Тамхана, T3 Даннета и C Даннета.

После того, как заданы все необходимые параметры, процедура One-Way ANOVA запускается нажатием на кнопку *OK*.

Ниже мы приводим результаты обработки данных предыдущего примера (см. табл. 10), чтобы у читателя была возможность сравнить выполнение процедуры однофакторного ДА в двух различных статистических системах.

Также, как и в системе Stadia, в результирующих таблицах даны очень полезные показатели описательной статистики (см. табл. 10). Очень удобно, что по каждой группе (уровню фактора) даны средние значения самооценок (столбец *Mean*) — это поможет исследователю правильно проинтерпретировать результаты парных сравнений, представленных ниже (см. табл. 13).

Таблица 10

Показатели описательной статистики

время решения задачи

	N	Среднее	Стд. отклонение	Стд. ошибка	95% доверительный интервал для среднего		Минимум	Максимум
					Нижняя граница	Верхняя граница		
Однородное	4	100.50	10.88	5.44	83.19	117.81	89	115
Неоднородное	4	97.50	16.42	8.21	72.37	123.63	80	112
Контрольное	4	66.00	12.19	6.10	46.60	85.40	49	77
Итого	12	88.00	20.30	5.86	75.10	100.90	49	115

Результаты теста Ливиня на гомогенность дисперсий выборок (1.413; p = 0.293) показывают, что одно из важных допущений ДА выполняется (см. табл. 11).

Таблица 11

Результаты проверки дисперсий на однородность

время решения задачи

Статистика Ливиня	ст.св.1	ст.св.2	Знч.
1.413	2	9	.293

Результаты, представленные в таблице 12, также показывают, что установлен статистически достоверный эффект фактора «условие решения задачи» на время ее решения ($F = 8.17$; $p = 0.009$). Кроме того, анализируя результаты парных сравнений средних, представленные в таблице 13, можно увидеть, что факторный эффект обусловлен главным образом различиями времени решения при контрольном условии и двух других условиях (34.5 и 31.5, соответственно). Стоит обратить внимание на то, что тест Геймса—Хауэлла не обнаружил достоверных различий между «неоднородным» и «контрольным» условиями, тогда как по Бонферрони они оказались статистически достоверными. Полученное несоответствие свидетельствует, с одной стороны, об известной относительности статистических выводов, а с другой — о необходимости осмысленно применять различные критерии адекватно имеющимся данным.

Таблица 12

Результирующая таблица ДА

время решения задачи

	Сумма квадратов	ст.св.	Средний квадрат	F	Знч.
Между группами	2922.000	2	1461.000	8.167	.009
Внутри групп	1610.000	9	178.889		
Итого	4532.000	11			

Таблица 13

Результаты множественного сравнения средних по трем экспериментальным условиям (тесты Бонферрони и Геймса—Хауэлла)

Зависимая переменная: время решения задачи

Текст	(I) Условие решения задачи	(J) Условие решения задачи	Средняя разность (I-J)	Стд. ошибка	Знч.	95% доверительный интервал	
						Нижняя граница	Верхняя граница
Бонферрони	Однородное	Неоднородное	3.00	9.458	1.000	-24.74	30.74
		Контрольное	34.50*	9.458	.016	6.76	62.24
	Неоднородное	Однородное	-3.00	9.458	1.000	-30.74	24.74
		Контрольное	31.50*	9.458	.026	3.76	59.24
	Контрольное	Однородное	-34.50*	9.458	.016	-62.24	-6.76
		Неоднородное	-31.50*	9.458	.026	-59.24	-3.76
Геймс-Хауэлл	Однородное	Неоднородное	3.00	9.458	.951	-28.60	34.60
		Контрольное	34.50*	9.458	.013	9.34	59.66
	Неоднородное	Однородное	-3.50	9.458	.951	-34.60	28.60
		Контрольное	31.50	9.458	.054	-.65	63.65
	Контрольное	Однородное	-34.50*	9.458	.013	-59.66	-9.34
		Неоднородное	-31.50	9.458	.054	-63.65	.65

*. Средняя разность значима на .05 уровне.

Задания

1. Определите число степеней свободы для межгрупповой ($\hat{S}^2_{BG}$) и внутригрупповой ($\hat{S}^2_{WG}$) дисперсий в следующих трех примерах:

№ примера	Число уровней фактора (групп), *k*	Число наблюдений в группе, *n*	$\hat{S}^2_{BG}$	$\hat{S}^2_{WG}$
1	3	10	?	?
2	4	20	?	?
3	2	30	?	?

2. Используя следующие числовые данные и табл. Приложения, рассчитайте *F*-отношение и оцените его значимость:

	Группа 1	Группа 2	Группа 3
n =	10	10	10
$\sum d^2_{total} = 382$ $\sum d^2_{BG} = 79$	$\hat{S}^2_{BG} = ?$ $\hat{S}^2_{WG} = ?$ *F* = ?		

3. Используя следующие числовые данные и табл. Приложения, рассчитайте *F*-отношение и оцените его значимость:

	Группа 1	Группа 2	Группа 3
n =	9	8	10
$\sum d^2_{total} = 192$ $\sum d^2_{BG} = 84$	$\hat{S}^2_{BG} = ?$ $\hat{S}^2_{WG} = ?$ *F* = ?		

4. С помощью одной из статистических систем проведите ДА данных, полученных в описанном ниже эксперименте.

Три группы испытуемых (опытные операторы-профессионалы, операторы-новички и студенты, не имевшие опыта операторской работы) выполняли задачу по слежению за движущимся объектом. По 10 опытам, проведенным с каждым испытуемым, было рассчитано среднее количество ошибок. Определите, зависит ли число ошибок от профессионального опыта испытуемых? Какие группы испытуемых значимо отличаются друг от друга по числу ошибок?

Опытные операторы	Операторы-новички	Студенты
3.13	1.39	5.47
3.25	5.38	5.6
3.64	4.07	6.88
3.4	3.87	6.4
2.59	4.37	3.02
1.97	3.79	6.18
3.16	3.33	5.52
4.22	5.39	4.15
1.36	3.37	2.07
3.47	4.74	4.68

5. Используя результаты вычислений, полученных в задании 3, оцените силу влияния фактора, рассчитав показатели η^2 (*эта квадрат*) и h^2 (Снедекора).

6. Используя данные из задания 4, проверьте гипотезу об однородности дисперсий трех выборок с помощью одной из статистических систем.

Глава 3

Двухфакторный дисперсионный анализ

3.1. Особенности двухфакторного ДА

Следует сразу же отметить, что принципиальной разницы между многофакторным и однофакторным ДА нет. Многофакторный анализ[1] не меняет общую логику ДА, а лишь несколько усложняет ее, поскольку, кроме учета влияния на зависимую переменную каждого из факторов по отдельности, следует оценивать и их совместное действие. Таким образом, то новое, что вносит в анализ данных многофакторный ДА, касается в основном возможности оценить *межфакторное взаимодействие*. Тем не менее, по-прежнему остается возможность оценивать влияние каждого фактора в отдельности. В этом смысле процедура многофакторного анализа (в варианте ее компьютерного использования) несомненно более экономична, поскольку всего за один запуск решает сразу две задачи: оценивает влияние каждого из факторов и их взаимодействие.

Общая схема двухфакторного эксперимента, данные которого обрабатываются ДА, представлена на рис. 2.

Данные, подвергаемые многофакторному ДА, часто обозначают в соответствии с количеством факторов и их уровней. Например, схема, представленная на рис. 2, называется двухфакторным планом 2×3. Если в исследовании было 3 фактора (A: 2 уровня, B: 3 уровня и C: 2 уровня), то схема получит

---

[1] В зарубежной литературе многофакторный анализ часто называется *дисперсионным факторным анализом* (Factorial Analysis of Variance), тем самым подчеркивается, что обрабатываются результаты многофакторных экспериментов. Отметим, что это название не следует путать с факторным анализом как самостоятельным методом многомерного анализа данных.

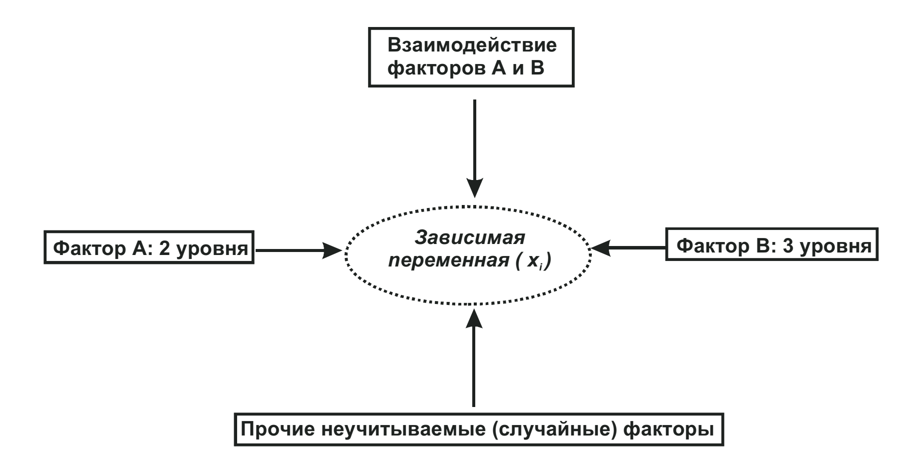


Рис. 2. Схема двухфакторного эксперимента

название 2×3×2. В зарубежной литературе для обозначения многофакторных схем также используют термин «n-way»: для двухфакторного — two-way, для трехфакторного — three-way и т.д. Таким образом, название «two-way 2×3 factorial analysis of variance» означает двухфакторный ДА с двумя уровнями первого фактора и тремя уровнями второго фактора.

В психологических исследованиях учет межфакторного взаимодействия может иметь особое значение и даже представлять для психолога основной интерес, поскольку несложно представить тот случай, когда каждый из факторов в отдельности может не оказывать на исследуемую характеристику никакого влияния, а сочетание двух или трех факторов оказывает существенное и закономерное влияние. Такие интересные феномены, когда совместное влияние факторов нельзя объяснить влиянием каждого фактора в отдельности, и означают взаимодействие факторов. Например, в возрастной психологии известно множество феноменов, когда проявление определенной психологической характеристики можно наблюдать только в определенном возрасте (фактор 1) и только в конкретном социальном окружении (фактор 2), т.е. при *определенном сочетании* указанных факторов. Вспомним классические эксперименты, подтверждающие известный

психологический закон Йеркса—Додсона: задачи разной сложности (фактор 1) решаются при различном уровне мотивации (фактор 2), и только *определенное сочетание* сложности и мотивации позволяет достичь оптимального уровня исполнения. В современной когнитивной психологии нередки случаи, когда исследователи находят и закономерные трехфакторные взаимодействия.

3.2. Линейная модель двухфакторного ДА

Линейная модель, лежащая в основе двухфакторного ДА, является простым распространением рассмотренной выше модели однофакторного ДА. Воспользуемся еще раз примером эксперимента по измерению ВР (см. гл. 1). Однако, примем во внимание, что в исследовании будет контролироваться еще один фактор — пол испытуемого. Тогда, если $\mu_{общ.}$ — среднее ВР в популяции взрослых людей, τ_j — дополнительный компонент, соответствующий влиянию возраста испытуемых (т.е. насколько среднее j-го возраста отличается от среднего генеральной совокупности: $\tau_j = \mu_j - \mu_{общ.}$), τ_g — компонент ВР, соответствующий влиянию пола испытуемого (т.е. насколько среднее ВР мужчин или женщин отличается от среднего генеральной совокупности: $\tau_g = \mu_g - \mu_{общ.}$), as_{jg} — вклад межфакторного взаимодействия (возраст×пол), ε_{ijg} — вклад уникальности конкретного (i-го) испытуемого («ошибка» линейной модели):

$$X_{ijg} = \mu_{общ.} + \tau_j + \tau_g + as_{jg} + \varepsilon_{ijg}. \qquad (26)$$

Как видно из этой простой формулы, межфакторное взаимодействие (as_{jg}) является неотъемлемой частью результата измерения независимой переменной. Иногда его влияние может быть достаточно сильным, чтобы «перевесить» влияние обоих изучаемых факторов. Например, представим, что изучается скорость чтения у детей разного возраста (фактор 1) в условиях разного социального окружения (фактор 2). Вполне вероятно, что некоторые дошкольники в силу присущей им стеснительности не смогут нормально выполнить предложенное им задание в окружении сверстников, т.е. опреде-

ленное взаимодействие факторов «возраст×социальное окружение» почти полностью подавляет продуктивную деятельность ребенка.

3.3. Проблема взаимодействия факторов

Рассмотрим проблему взаимодействия факторов подробнее. Когда мы оцениваем взаимодействие двух факторов, то фактически нас интересует, будет ли одинаковым влияние на зависимую переменную одного из факторов *на всех уровнях* другого фактора. Очевидно, что, если это влияние неодинаково, то второй фактор каким-то образом опосредует влияние первого фактора и, следовательно, можно говорить о существовании взаимодействия между ними.

Используя один вымышленный эксперимент по изучению словарного запаса у младших школьников, рассмотрим различные варианты возможных результатов. Пусть в нашем эксперименте школьникам 2-го класса, мальчикам и девочкам (*фактор 1*: пол ребенка), было предложено составить список знакомых им слов, относящимся к трем различным сферам жизни человека: домашнее хозяйство, природа и спорт (*фактор 2*: сфера жизни). Предполагается, что в силу специфики интересов у детей данного возраста мальчики и девочки могут обладать различным словарным запасом, относящимся к разным сферам жизни.

Пример 1, результаты которого представлены на рис. 3, демонстрирует четкую зависимость количества воспроизведенных слов от пола ребенка и заданной в качестве образца сферы жизни. В силу особенностей интересов детей этой возрастной группы словарный запас мальчиков и девочек оказался различным в разных сферах жизни — налицо эффект взаимодействия обоих контролируемых факторов.

Пример 2 показывает другой возможный вариант результатов эксперимента, где также было обнаружено взаимодействие факторов (см. рис. 4). Эти результаты показывают, что словарный запас у девочек больше, чем у мальчиков, но это преимущество заметно отличается для различных сфер жизни. Таким образом, здесь также мы можем обнаружить взаимодействие двух факторов.

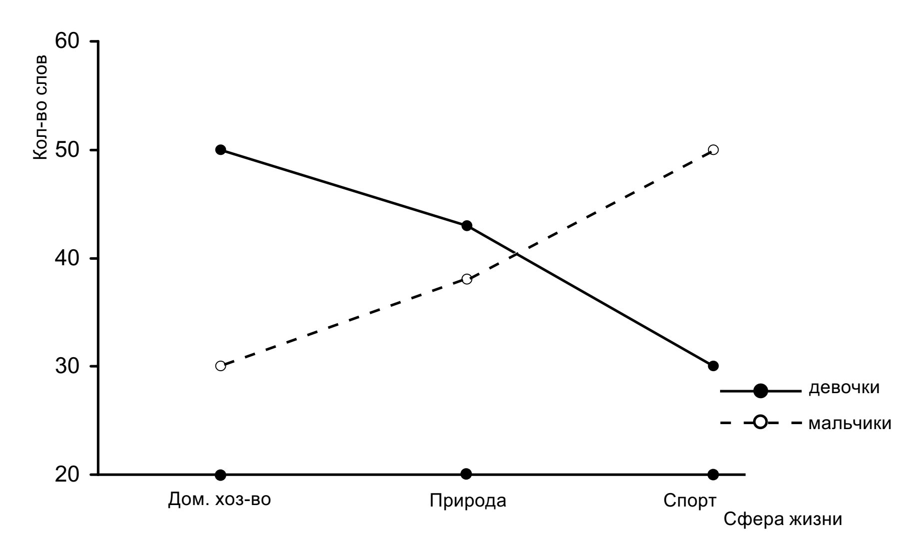


Рис. 3. Результаты двухфакторного эксперимента по изучению словарного запаса у младших школьников: имеет место взаимодействие факторов «пол×сфера жизни»

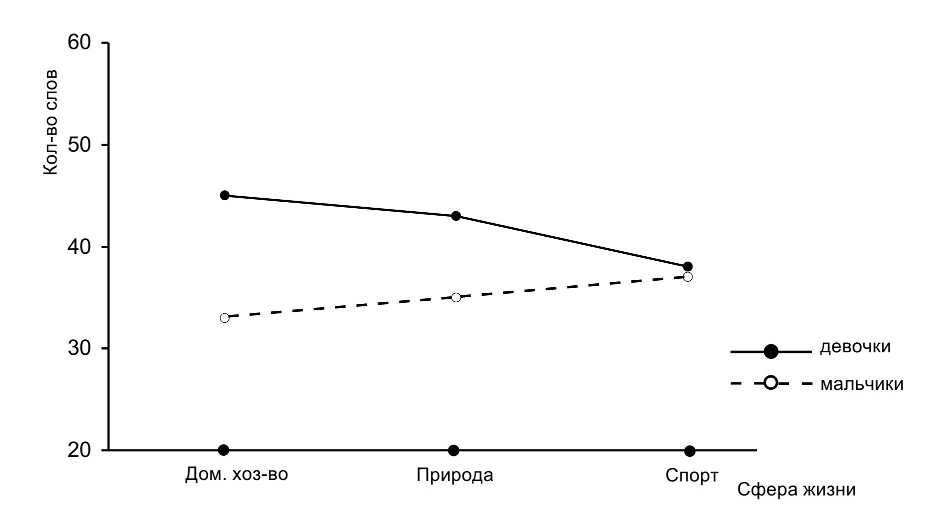


Рис. 4. Результаты двухфакторного эксперимента по изучению словарного запаса у младших школьников: имеет место взаимодействие факторов «пол×сфера жизни»

И, наконец, ***пример 3*** демонстрирует отсутствие межфакторного взаимодействия (см. рис. 5). Фактическая парал-

лельность сплошной и пунктирной линий свидетельствует о том, что словарный запас у девочек больше, чем у мальчиков, причем это преимущество *одинаково* во всех сферах. Это и означает, что в данном примере влияние фактора «пол» на словарный запас *не зависит* от того, в какой сфере жизни он оценивается. Таким образом, взаимодействие факторов «пол×сфера жизни» не обнаруживается.

Графический анализ результатов эксперимента на уровне средних арифметических является хорошим средством для формулирования предварительных гипотез, однако сам факт «параллельности» (см. рис. 5) или «перекрещенности» (см. рис. 3) двух кривых, естественно, не дает возможности достоверно и определенно говорить об установлении факта взаимодействия двух факторов. Нередко по графику можно переоценить или же недооценить факт межфакторного взаимодействия. Однако после проведения ДА бывает очень полезно обратиться к графическому изображению анализируемых зависимостей, чтобы более полно понять суть обнаруженных влияний факторов и/или межфакторного взаимодействия.

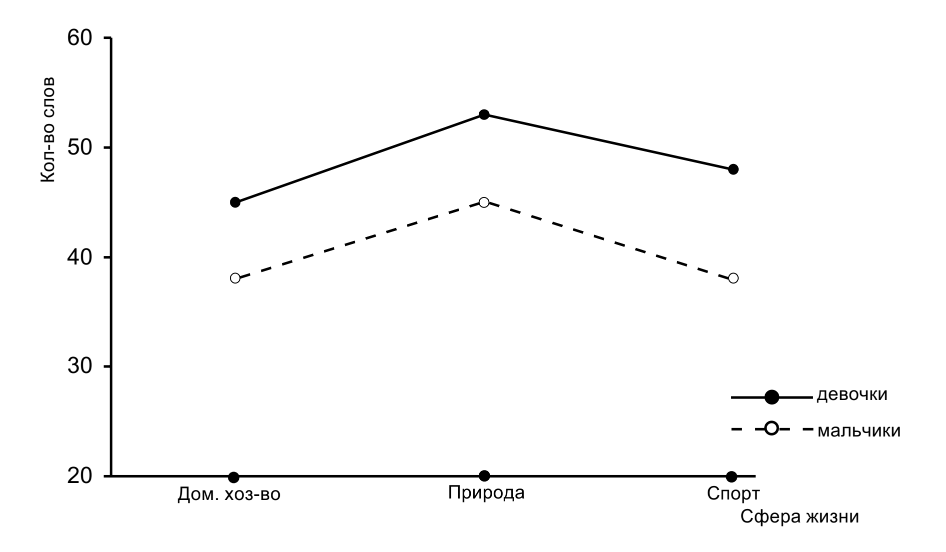


Рис. 5. Результаты двухфакторного эксперимента по изучению словарного запаса у младших школьников: отсутствует взаимодействие факторов «пол×сфера жизни»

3.4. Обработка данных двухфакторного эксперимента

Рассмотрим конкретный числовой пример с данными двухфакторного эксперимента и проведем все необходимые для выполнения двухфакторного ДА вычисления[1]. Воспользуемся данными эксперимента по исследованию зависимости кратковременной памяти от темпа предъявления слов и их длины (см. табл. 14). В этом эксперименте участвовало 6 различных групп испытуемых по 5 человек в каждой. С каждым испытуемым проводилось по 3 повторных опыта и их данные усреднялись. Первый контролируемый фактор (фактор А) — длина тестового слова — включал 3 уровня: 3, 5 и 7 букв. Второй фактор (фактор В) — темп предъявления слов — включал 2 уровня: низкий темп и высокий темп. До начала выполнения вычислений проделаем небольшой визуальный анализ полученных результатов (см. рис. 6). Представленные на рис. 6 результаты позволя-

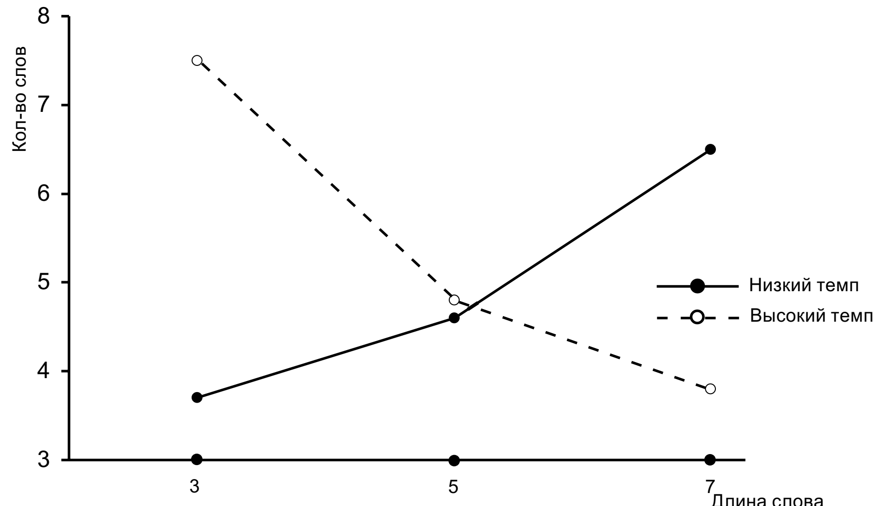


Рис. 6. Зависимость объема памяти от длины тестового слова и темпа его предъявления:
ось X — длина тестового слова; ось Y — количество запомненных слов

---

[1] Мы специально взяли пример с равным количеством наблюдений (испытуемых) для каждого уровня обоих факторов для удобства и простоты вычислений. Пример с неравным количеством наблюдений рассматриваться не будет, поскольку нам важно показать лишь общую логику вычислений в ДА, и мы надеемся, что современному психологу не придется что-то делать вручную, а не с помощью какой-либо статистической системы.

Таблица 14

Результаты эксперимента по исследованию кратковременной памяти

№№ строк	*Фактор А — длина тестового слова*					
	3 буквы		5 букв		7 букв	
Фактор В — темп пред-я	Низкий	Высокий	Низкий	Высокий	Низкий	Высокий
1	4.8	7.5	5.3	6.0	7.2	3.8
2	3.2	7.8	4.1	4.2	6.8	3.2
3	3.1	9.2	3.9	3.8	5.4	4.8
4	3.9	7.3	4.7	4.8	5.9	3.7
5	3.8	5.8	5.0	5.2	6.3	3.3
6 По группам (j)	$n_1 = 5$ $\Sigma_1 = 18.8$	$n_2 = 5$ $\Sigma_2 = 37.6$	$n_3 = 5$ $\Sigma_3 = 23.0$	$n_4 = 5$ $\Sigma_4 = 24.0$	$n_5 = 5$ $\Sigma_5 = 31.6$	$n_6 = 5$ $\Sigma_6 = 18.8$
7 По фактору А	$n_A = 10$ $\Sigma_{A1} = 56.4$		$n_A = 10$ $\Sigma_{A2} = 47.0$		$n_A = 10$ $\Sigma_{A3} = 50.4$	
8 По фактору В	$n_B = 15$ $\Sigma_{B1} = 73.4$			$n_B = 15$ $\Sigma_{B2} = 80.4$		
9 В целом	$N_{total} = 30$, $\Sigma_{total} = 153.8$, $s^2_{total} = 2.49$.					
10 Пояснения к таблице	Строки 1—5: число правильно запомненных слов в отдельных группах. $n_{1—6}$ — число испытуемых в отдельных группах. $\Sigma_{1—6}$ — суммы правильно запомненных слов в отдельных группах. $\Sigma_{A1—A3,\ B1—B3}$ — суммы правильно запомненных слов, соответствующие каждому уровню каждого фактора. N_{total}, Σ_{total}, и s^2_{total} — число испытуемых, сумма всех правильно запомненных слов и их дисперсия, соответственно, по группе в целом.					

ют предположить, что имела место зависимость успешности запоминания слов как от их длины, так и от скорости предъявления на экране. Более того, можно ожидать наличия эффекта взаимодействия этих факторов, т.е. при высоком темпе предъявления лучше запоминались короткие слова, тогда как при низком — длинные. Попробуем, используя двухфакторный ДА, количественно проверить наши предположения.

До начала вычислений сформулируем *статистические гипотезы:*

1. **Отдельно по фактору А:**

• H_0: количество запомненных слов не зависит от длины слова, т.е. вариации $\overline{X}_3$, $\overline{X}_5$ и $\overline{X}_7$ — случайны.

• H_1: количество запомненных слов зависит от длины слова, т.е. вариации $\overline{X}_3$, $\overline{X}_5$ и $\overline{X}_7$ обусловлены влиянием фактора А.

2. **Отдельно по фактору В:**

• H_0: количество запомненных слов не зависит от темпа предъявления слова, т.е. вариации $\overline{X}_н$ и $\overline{X}_в$ — случайны.

• H_1: количество запомненных слов зависит от темпа предъявления слова, т.е. вариации $\overline{X}_н$ и $\overline{X}_в$ обусловлены влиянием фактора В.

3. **Для взаимодействия факторов А и В:**

• H_0: влияние фактора А на количество запомненных слов одинаково при разных темпах предъявления слова, и наоборот.

• H_1: влияние фактора А на количество запомненных слов различно при разных темпах предъявления слова, и наоборот.

Сначала, как и в предыдущей части, рассчитаем *общую* по всем испытуемым сумму квадратов:

$$\sum d^2_{total} = (N-1)\cdot s^2_{total} = 29\cdot 2.49 = 72.21.$$

Далее оценим *межгрупповую (или общую факторную)* сумму квадратов:

$$\sum d^2_{BG} = \sum_{j=1}^{k}\frac{(\sum X_i)^2}{n} - \frac{(\sum_{i=1}^{N} X_i)^2}{N} =$$

$$= \frac{18.8^2+37.6^2+23.0^2+24.0^2+31.6^2+18.8^2}{5} - \qquad (27)$$

$$-\frac{(153.8)^2}{30} = 844.84 - 788.48 = 56.36,$$

где n — число наблюдений в каждой группе, j — знак суммирования средних по k группам, а i — индекс суммирования значений всех наблюдений в отдельности. Фактически межгрупповая, или общая факторная сумма квадратов является оценкой вклада линейной двухфакторной модели (которую мы и используем!) в общую сумму квадратов. Поэтому ее иногда называют *модельной*.

Внутригрупповую (или остаточную) сумму квадратов вычисляем как разность между общей факторной и межгрупповой суммами квадратов:

$$\sum d_{WG}^2 = \sum d_{total}^2 - \sum d_{BG}^2 = 72.21 - 56.36 = 15.85. \quad (28)$$

Далее определяем *суммы квадратов отклонений для факторов A и B*. При расчете $\sum d_A^2$ мы усредняем данные по уровням фактора B, а при расчете $\sum d_B^2$ — по уровням фактора A.

$$\sum d_A^2 = \sum_1^a \frac{\sum (X_A)^2}{n \cdot b} - \frac{(\sum_{i=1}^{N} X_i)^2}{N} =$$

$$\frac{(56.4^2 + 47.0^2 + 50.4^2)}{5 \cdot 2} - 788.48 = \quad (29)$$

$$= 793.01 - 788.48 = 4.53,$$

где b — число уровней фактора B, X_A — наблюдения, сделанные по каждому уровню фактора A и усредненные по фактору B.

$$\sum d_B^2 = \sum_1^b \frac{\sum (X_B)^2}{n \cdot a} - \frac{(\sum_{i=1}^{N} X_i)^2}{N} =$$

$$= \frac{(73.4^2 + 80.4^2)}{5 \cdot 3} - 788.48 = \quad (30)$$

$$= 790.11 - 788.48 = 1.63,$$

где a — число уровней фактора A, X_B — наблюдения, сделанные по каждому уровню фактора B и усредненные по фактору A.

Поскольку сумма квадратов межфакторного взаимодействия входит третьей составляющей в межгрупповую (общую факторную) сумму квадратов наряду с вкладом каждого из факторов, то мы ее вычисляем просто как разность:

$$\sum d_{AB}^2 = \sum d_{BG}^2 - (\sum d_A^2 + \sum d_B^2) = \\ = 56.36 - (4.53 + 1.63) = 50.20. \quad (31)$$

Далее переходим к установлению числа степеней свободы для каждой суммы квадратов:

1. Для общей: $df_{total} = N - 1 = 30 - 1 = 29.$

2. По фактору А: $df_A = a - 1 = 3 - 1 = 2.$

3. По фактору В: $df_B = b - 1 = 2 - 1 = 1.$

4. Для взаимодействия факторов А и В: $df_{AB} = df_A \times df_B = 2 \times 1 = 2.$

5. Для внутригрупповой или остаточной: $df_{WG} = N - a \times b = 30 - 3 \times 2 = 30 - 6 = 24.$

6. Для межгрупповой, или общей факторной (или модельной): $df_{BG} = a \times b - 1 = 3 \times 2 - 1 = 5.$

Затем находим оценки дисперсий, входящих в F-отношения, поделив суммы квадратов на соответствующие значения степеней свободы:

1. $\hat{S}_A^2 = \dfrac{4.53}{2} = 2.27.$

2. $\hat{S}_B^2 = \dfrac{1.63}{1} = 1.63.$

3. $\hat{S}_{AB}^2 = \dfrac{50.20}{2} = 25.1.$

4. $\hat{S}_{WG}^2 = \dfrac{15.85}{24} = 0.66.$

5. $\hat{S}_{BG}^2 = \dfrac{56.36}{5} = 11.27.$

И наконец, вычисляем *F*-отношения и оцениваем их значимость, сравнивая с критическими из таблицы Приложения[1]:

1. $F_A = \dfrac{2.27}{0.66} = 3.44.$ $F_A > F_{кр.} = 3.40$, следовательно, на уровне значимости $p < 0.05$ мы отвергаем нулевую гипотезу и делаем вывод о статистической значимости влияния фактора А. Установив значимость фактора А, мы обнаружили один из двух *главных эффектов* [2], т.е. мы говорим о влиянии фактора «длина слова» как такового, игнорируя все остальные условия эксперимента.

2. $F_B = \dfrac{1.63}{0.66} = 1.97.$ $F_A < F_{кр.} = 4.26$, следовательно, на уровне значимости $p = 0.05$ у нас нет оснований отвергать нулевую гипотезу и мы делаем вывод о статистической незначимости влияния фактора В. *Главный эффект* фактора «темп предъявления» не обнаружен.

3. $F_{AB} = \dfrac{25.1}{0.66} = 38.03.$ $F_A > F_{кр.} = 3.40$, следовательно, на уровне значимости $p < 0.05$ мы отвергаем нулевую гипотезу и делаем вывод о статистической значимости влияния взаимодействия факторов А и В.

4. $F_{BG} = \dfrac{11.27}{0.66} = 17.08.$

---

[1] При расчете *F*-отношения следует учитывать, какие выборки данных по группам испытуемых были получены в эксперименте — связанные или несвязанные. В данном случае в разных группах были разные испытуемые, следовательно несвязанные выборки. В зарубежной литературе экспериментальные планы, в которых одни и те же испытуемые повторно участвуют в разных экспериментальных условиях, называются планами с повторными измерениями (repeated-measures designs).

[2] В отличие от термина *главный эффект,* когда мы оцениваем влияние одного фактора вне зависимости от уровней другого фактора, также используют термин *простой эффект,* когда мы рассматриваем влияние одного фактора только на одном из уровней другого. В нашем примере — это рассмотреть зависимость количества запомненных слов от длины слова только при высоком темпе предъявления слов.

Как следует из результатов ДА (см. табл. 15), мы обнаружили один главный эффект: зависимость успешности запоминания от длины запоминаемого слова. Обнаружить зависимость успешности запоминания от темпа предъявления не удалось. Однако с высокой достоверностью был обнаружен эффект взаимодействия исследуемых факторов, т.е. количество запоминаемых слов зависит от сочетания длины слова и темпа его предъявления. Кроме того, мы обнаружили значимое соответствие двухфакторной модели имеющимся данным, о чем свидетельствует высокая достоверность вклада общей факторной компоненты в общую дисперсию оценок. Поэтому модель ДА, использованная нами для анализа данного примера, хорошо соответствует полученным данным. Хорошей мерой согласия модели и данных может быть отношение суммы квадратов, предсказанных моделью, к общей сумме квадратов: $\frac{56{,}4}{72.12} = 0.78$. В нашем случае модель объясняет 78% общей суммы квадратов. В зависимости от того, много это или мало, исследователь может добавлять в модель дополнительные факторы или ограничиться имеющимися.

Таблица 15

Результаты двухфакторного ДА

Источник вариации		Степени свободы, *df*	$\sum d^2$	$\hat{s}^2$	*F*	*P*
Главные эффекты	По фактору A	2	4.53	2.27	**3.44**	**<0.05**
	По фактору B	1	1.63	1.63	1.97	>0.05
	Взаимодействие A × B	2	50.20	2.51	**38.03**	**<0.01**
	Общая факторная (модель)	5	56.36	11.27	**17.08**	**<0.01**
	Остаточная	24	15.85	0.66	-	-
	Общая	29	72.21	2.49	-	-

3.5. Множественные сравнения средних в двухфакторном ДА

Как и в случае однофакторного ДА, методы множественного сравнения средних также необходимы, чтобы конкрет-

но выяснить, какие пары выборочных средних, относящихся к различным уровням обоих факторов, отличаются друг от друга. Все соответствующие рассуждения предыдущего раздела безусловно применимы и к двухфакторному варианту ДА. Тем не менее, в формулах есть небольшое различие. При необходимости выполнить вычисления вручную, а не с помощью одной из статистических систем, стоит обратить особое внимание на интерпретацию «*n*» в знаменателе: это число определяется как произведение числа наблюдений в отдельной группе (n_i) на число уровней данного фактора (a_i или b_i).

3.6. Оценка величины факторного эффекта

Как и в предыдущем разделе, в качестве показателей силы факторного влияния могут быть рекомендованы индексы η^2 и ω^2. Первый индекс рассчитывается аналогичным образом: оценивается отношение соответствующей суммы квадратов к общей сумме квадратов. Для расчета более сложного и адекватного индекса ω^2 необходимы дополнительные вычисления. Для эксперимента с фиксированными эффектами (как в нашем примере) это делается следующим образом:

$$\omega^2_{A,B,AB} = \frac{\hat{\sigma}^2_{A,B,AB}}{\hat{\sigma}^2_{total}} = \frac{\hat{\sigma}_{A,B,AB}}{\hat{\sigma}^2_A + \hat{\sigma}^2_B + \hat{\sigma}^2_{AB} + \hat{\sigma}^2_e}, \qquad (32)$$

т.е. находится отношение *скорректированной* оценки дисперсии каждого фактора или их взаимодействия к скорректированной оценке общей дисперсии.

Скорректированные оценки дисперсий рассчитываются так:

$$\hat{\sigma}^2_A = \frac{(a-1)(\hat{S}^2_A - \hat{S}^2_{WG})}{n \cdot a \cdot b}, \qquad (33)$$

$$\hat{\sigma}^2_B = \frac{(b-1)(\hat{S}^2_B - \hat{S}^2_{WG})}{n \cdot a \cdot b}, \qquad (34)$$

$$\hat{\sigma}^2_{AB} = \frac{(a-1)(b-1)(\hat{S}^2_{AB} - \hat{S}^2_{WG})}{n \cdot a \cdot b}, \tag{35}$$

$$\hat{\sigma}^2_e = \hat{S}^2_{WG}. \tag{36}$$

Сделаем соответствующие расчеты для нашего примера:

1. $\omega^2_A = \dfrac{\hat{\sigma}^2_A}{\hat{\sigma}^2_{total}} = \dfrac{0.107}{0.922} = 0.12.$

2. $\omega^2_B = \dfrac{\hat{\sigma}^2_B}{\hat{\sigma}^2_{total}} = \dfrac{0.032}{0.922} = 0.03.$

3. $\omega^2_{AB} = \dfrac{\hat{\sigma}^2_{AB}}{\hat{\sigma}^2_{total}} = \dfrac{0.123}{0.922} = 0.13.$

Таким образом, сила влияния фактора A составляет 12%, фактора B — 3%, а их межфакторного взаимодействия — 13%.

Для моделей со случайными или смешанными эффектами (когда один или несколько факторов имеют фиксированное количество уровней, а другой/другие — случайное) формулы для расчета скорректированных оценок дисперсий несколько модифицируются (см., напр., *Хауэлл*, 1998).

3.7. Оценка простых эффектов

Простым эффектом фактора в отличие от *основного эффекта* называется эффект одного фактора, рассматриваемый только на одном из уровней другого фактора. Анализ простых факторных эффектов бывает очень полезным для более тщательного и подробного анализа причин уже установленного в ходе ДА межфакторного взаимодействия, когда исследователь хочет ответить на вопрос, в чем же выражается конкретно его эффект. Поскольку он может проявиться в большей или меньшей степени лишь на одном из уровней другого фактора, то вполне резонно посмотреть, на каком из них это взаимодействие имеет место и в какой степени. Ту же задачу выполняет и использование множественного сравнения средних.

Рассмотрим на нашем примере, как проявляется взаимодействие факторов А (длина слова) и В (темп предъявления) раздельно на каждом из уровней фактора В (см. табл. 16).

Таблица 16

Суммы по столбцам, взятые отдельно для каждого уровня факторов А из табл. 14.

Кол-во букв / Темп	3 буквы (A_1)	5 букв (A_2)	7 букв (A_3)	Всего по фактору А
Низкий (B_1)	18.8	23.0	31.6	**73.4**
Высокий (B_2)	37.6	24.0	18.8	**80.4**
Кол-во наблюдений (испытуемых)	5	5	5	15×2=30

Оценим сумму квадратов по фактору А для *первого уровня* фактора В:

$$\sum d^2_{A-B_1} = \frac{(18.8^2 + 23.0^2 + 31.6^2)}{5} - \frac{73.4^2}{15} = 375.2 - 359.17 = 17.03.$$

Оценим сумму квадратов по фактору А для *второго уровня* фактора В:

$$\sum d^2_{A-B_2} = \frac{(37.6^2 + 24.0^2 + 18.8^2)}{5} - \frac{80.4^2}{15} =$$

$$= 468.64 - 430.94 = 37.70$$

Далее рассчитаем все остальные показатели и сведем их стандартную таблицу (см. табл. 17).

Таблица 17

Результаты ДА простых эффектов для фактора А

Источник вариации	Степени свободы, *df*	$\sum d^2$	$\hat{S}^2$	*F*	*P*
Фактор А для условия B_1	2	17.03	8.52	**12.91**	**<0.01**
Фактор А для условия B_2	2	37.70	18.85	**28.56**	**<0.01**
Остаточная	24	15.85	0.66	-	-

Как следует из результатов, представленных в таблице 17, влияние фактора А оказалось достоверно значимым на обоих уровнях фактора В, т.е. длина слова влияет на запоминание и при высоком, и при низком темпе предъявления слов.

Рассмотрим также влияние фактора В при трех различных уровнях фактора А (см. табл. 18).

Таблица 18

Суммы по столбцам, взятые отдельно для двух уровней фактора В из табл. 14

Темп / Кол-во букв	Низкий (B_1)	Высокий (B_2)	Всего
3 буквы (A_1)	18.8	37.6	**56.4**
5 букв (A_2)	23.0	24.0	**47.0**
7 букв (A_3)	31.6	18.8	**50.4**
Кол-во наблюдений (испытуемых)	5	5	10×3=30

Оценим сумму квадратов по фактору В для *первого уровня* фактора А:

$$\sum d^2_{B-A_1} = \frac{(18.8^2 + 37.6^2)}{5} - \frac{56.4^2}{10} = 353.44 - 318.1 = 35.44 .$$

Оценим сумму квадратов по фактору В для *второго уровня* фактора А:

$$\sum d^2_{B-A_2} = \frac{(23.0^2 + 24.0^2)}{5} - \frac{47.0^2}{10} = 221.0 - 220.9 = 0.1 .$$

Оценим сумму квадратов по фактору В для *третьего уровня* фактора А:

$$\sum d^2_{B-A_3} = \frac{(31.6^2 + 18.8^2)}{5} - \frac{50.4^2}{15} = 270.4 - 254.02 = 16.38 .$$

Далее рассчитаем все остальные показатели и сведем их стандартную таблицу (см. табл. 19).

Результаты, представленные в таблице 19, также показывают высокую достоверность влияния фактора В на первом и третьем уровнях фактора А, т.е. при условии предъявления ко-

Таблица 19

Результаты ДА простых эффектов для фактора В

Источник вариации	Степени свободы, *df*	$\sum d^2$	$\hat{S}^2$	*F*	*P*
Фактор В для условия A_1	1	35.44	35.44	**53.7**	**<0.01**
Фактор В для условия A_2	1	0.1	0.1	0.15	>0.05
Фактор В для условия A_3	1	16.38	16.38	**24.82**	**<0.01**
Остаточная	24	15.85	0.66	-	-

ротких и длинных слов темп предъявления оказывает значимое влияние на запоминание слов. В том же случае, когда использовались слова средней длины (5 букв), темп предъявления слов никак не влиял на их запоминание. Это достаточно интересный результат сам по себе, поскольку мы не обнаружили основного эффекта фактора В по эксперименту в целом, а, разложив его на составляющие, нашли значимые влияния. Результаты анализа простых эффектов фактора В могут в данном случае заставить исследователя задуматься о возможных психологических механизмах кратковременной памяти, тогда как простая формальная констатация отсутствия основного эффекта фактора В может «замаскировать» имеющиеся результаты.

Упомянув термин «разложение на составляющие», уточним, из каких «составляющих» состоит сумма квадратов простого эффекта одного из факторов в многофакторном эксперименте. Например, сумма квадратов простых эффектов фактора А (по уровням фактора В) равна сумме квадратов самого фактора А и межфакторного взаимодействия А×В:

$$\sum d^2_{A-B_1} + \sum d^2_{A-B_2} = \sum d^2_A + \sum d^2_{AB} \,. \qquad (37)$$

Таким образом, сумма простых эффектов фактора соответствует совокупности всех его влияний, т.е. сумме основного эффекта и межфакторного взаимодействия[1].

Аналогично оценкам простых эффектов каждого из факторов в отдельности, можно оценить *простой эффект взаимодействия факторов*. Это бывает полезно, когда, например,

[1] Проделав простые арифметические вычисления, читатель может сам убедиться в справедливости этой формулы.

результаты *трехфакторного* ДА не обнаружили статистически достоверного эффекта межфакторного взаимодействия факторов А×В, но исследователь ожидает наличие такового на одном из уровней третьего фактора С.

Вычисления, которые мы в данном случае опустим из-за их простоты и очевидности, в целом аналогичны проделанным выше. При необходимости читатель может обратиться к книге Хауэлла (1998), где подробно разобран соответствующий пример.

3.8. Двухфакторный ДА с выборками неравного размера

Хотя чаще всего исследователь планирует эксперимент с выборками равного размера ($n_1 = n_2 = n_3 = ... = n_i$), в силу ряда причин это не всегда осуществимо: испытуемые могут заболеть или вообще отказаться от участия в опытах и т.п. Если в случае с однофакторным ДА проблем не возникает, то с многофакторным ДА имеют место сложности. Конечно, как рекомендуется в ряде руководств, можно выровнять количество испытуемых в группах, отбросив случайным образом «лишних» из каждой группы. Однако, иногда это неразумно или вообще невозможно в силу исходной малочисленности выборок.

Сложности возникают в связи с тем, что при выборках неравного размера строки, столбцы и эффекты межфакторного взаимодействия уже *не являются независимыми*. Не вдаваясь глубоко в эту проблему, ограничимся лишь ее обозначением. На наш взгляд, наилучшим выходом будет использование одной из современных статистических систем, где эта проблема находит свое решение. Однако, если у читателя возникнет желание разобраться в способах расчетов, мы можем отослать его к вполне доступной для психолога литературе (см. *Лакин*, 1980; *Хауэлл*, 1998; *Гласс*, *Стенли*, 1976).

3.9. Сложности ДА с большим количеством факторов

Все основные принципиальные положения, касающиеся двухфакторного ДА, безусловно применимы и к обработке

трех-, четырех- и более факторных экспериментов. Тем не менее, сложность вычислений и вероятность появления в них ошибки значительно возрастают, поэтому мы рекомендуем коллегам-психологам использовать для реализации ДА одну из современных компьютерных статистических систем.

Еще несколько слов посвятим некоторым важным терминам, использующимся в статистических системах, и описаниях к ним. Когда мы обрабатываем результаты двухфакторного эксперимента или рассматриваем двухфакторные взаимодействия в многофакторном эксперименте, то говорят о взаимодействии двух факторов (A×B, A×C или B×C) или *взаимодействии первого порядка*. Когда же мы рассматриваем взаимодействие сразу трех факторов (A×B×C), то его называют *взаимодействием второго порядка*.

Содержательная интерпретация межфакторных взаимодействий более чем первого порядка может представлять для исследователя значительную проблему. Рассмотрим несколько формальных соображений по поводу трехфакторного взаимодействия A×B×C. Одним из вариантов объяснения этого взаимодействия может быть мысль о том, что взаимодействие факторов A×B само по себе взаимодействует с фактором C. Рассмотрим эту типичную ситуацию более подробно графически (см. рис. 7). Обратите внимание на то, что картина взаимодействия A×B на уровне C_1 (левая часть рисунка) принципиально отличается от такового на уровне C_2 (правая часть рисунка) — фактически зависимость A от B изменилась на противоположную. Это и означает, что имеет место трехфакторное взаимодействие — A×B×C. Но та же самая картина будет иметь место при взаимодействии A×C на различных уровнях B или B×C на различных уровнях A. Кроме того, можно предположить, что один из факторов A или B сам взаимодействует с фактором C, и тогда все будет выглядеть еще сложнее. Поэтому с формальной точки зрения трехфакторные взаимодействия достаточно непросто и неоднозначно интерпретируемы, и психологу требуются четкие содержательные гипотезы, чтобы сделать какой-либо определенный вывод.

Казалось бы, что если в результате трехфакторного ДА получен всего один значимый основной эффект, то его ин-

терпретация проста: значит данный фактор влияет на независимую переменную. Однако эта простота ошибочна и мы не можем сделать однозначного и простого вывода о влиянии основного эффекта этого фактора, поскольку этот эффект зависит от двух других факторов тоже. Поэтому, даже при высокой значимости основного эффекта одного фактора, мы не можем только на основании этого результата идентифицировать источник различий между группами. В случае, когда мы обнаружили 1—2 основных эффекта и 1—2 межфакторных взаимодействия, нужно быть особенно осторожными и не делать поспешных выводов из формальных результатов ДА. Стоит особенно тщательно анализировать графики, показывающие взаимоотношение факторов и соотносить их с результатами парных сравнений средних.

При планировании реальных экспериментов следует с особой осторожностью включать в экспериментальный план взаимодействия высших порядков и без особой необходимости не делать этого, поскольку post factum всегда возникает проблема интерпретации обнаруженных взаимодействий.

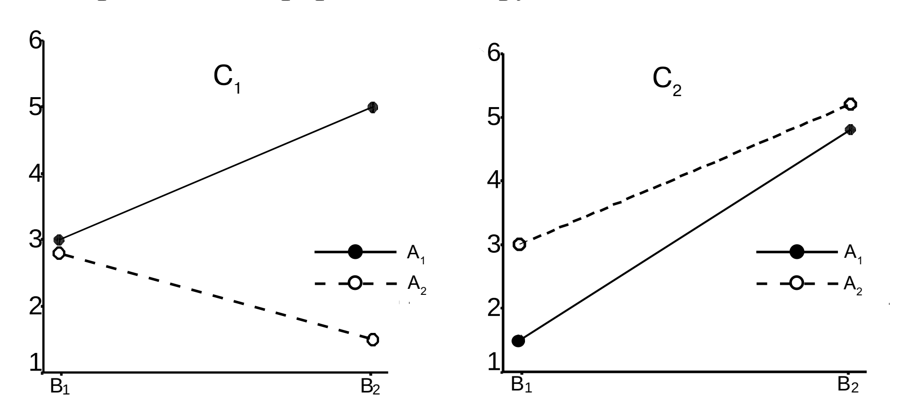


Рис. 7. Межфакторные взаимодействия второго порядка: проявление взаимодействия факторов A×B при двух различных уровнях фактора C:

C_1 — слева; C_2 — справа; A_1, A_2, B_1 и B_2 — выраженность признака при двух различных уровнях факторов A и B

Контрольные вопросы

1. Что такое многофакторный ДА, чем он отличается от однофакторного ДА?

2. В чем заключается межфакторное взаимодействие? Как показать графически наличие и отсутствие межфакторного взаимодействия?

3. Что стоит за обозначением «трехфакторный план 3×3×2»?

4. Объясните, что означает английское название «two-way 2×3 factorial analysis of variance».

5. Какая математическая модель лежит в основе многофакторного ДА? Напишите формулу двухфакторной модели и объясните ее.

6. Сформулируйте статистические гипотезы (H_0 и H_1) для двухфакторного ДА.

7. Что такое общая сумма квадратов и как она рассчитывается?

8. Что стоит за межгрупповой (общей факторной) суммой квадратов и как она рассчитывается?

9. Какое значение имеет внутригрупповая (остаточная) сумма квадратов, как она рассчитывается?

10. Как рассчитываются суммы квадратов для факторов в двухфакторном ДА?

11. Как связаны между собой суммы квадратов факторов и сумма квадратов межфакторного взаимодействия?

12. Как оценить степени свободы для межфакторного взаимодействия и внутригрупповой дисперсии?

13. Что такое главный и простой эффекты фактора?

14. С помощью каких индексов оценивают силу факторных эффектов и межфакторного взаимодействия?

15. Для чего бывает полезно оценить простой факторный эффект или простой эффект факторного взаимодействия?

16. Что такое взаимодействие второго порядка?

17. В чем сложности интерпретации результатов многофакторных взаимодействий?

3.10. Работа с двухфакторным дисперсионным анализом в статистической системе STADIA 6.0

В системе STADIA реализованы два варианта двухфакторного ДА — с повторными измерениями и без повторных измерений. В соответствии с используемой моделью (т.е. имеющимися данными) данные в электронную таблицу заносятся по-разному. В случае более чем двух факторов предусмотрена отдельная процедура обработки данных, однако в ней не учитываются межфакторные взаимодействия.

Вариант ДА без повторных измерений. В эксперименте без повторных измерений каждому сочетанию факторов соответствует *всего одно измерение* (а не *n*, как было в нашем примере)[1]. Поэтому исходные данные заносятся в электронную таблицу в виде матрицы размером $n \times m$, столбцы которой должны соответствовать различным уровням первого фактора, строки — различным уровням второго фактора, а каждая ячейка, соответственно, содержит только одно измерение.

Рассмотрим ***конкретный пример***. Двухфакторный план 5×5. В эксперименте изучалась зависимость эффективности обучения иностранному языку от методики обучения (фактор А: 5 уровней) и возраста обучаемых (фактор В: 5 уровней). В электронную таблицу заносятся числа, соответствующие количеству ошибок, допущенных каждым обучающимся (в каждой из 5 групп по одному человеку) по итогам окончательного тестирования полученных знаний (см. табл. 20).

Таблица 20

Результаты психолого-педагогического эксперимента по эффективности обучения иностранному языку

Возраст/Методика	Мет-ка 1	Мет-ка 2	Мет-ка 3	Мет-ка 4	Мет-ка 5
Возраст 1	60	90	60	20	60
Возраст 2	40	70	80	30	50
Возраст 3	90	30	100	70	40
Возраст 4	80	40	140	40	100
Возраст 5	120	110	130	90	140

[1] Как правило, вариант ДА без повторных измерений используют, когда число уровней обоих факторов достаточно велико (4—5). В противном случае интуитивно ясно, что никакая статистика при недостатке данных «не сработает».

После ввода данных и проверки его правильности переходим в меню статистических методов (**F9**). Там следует выбрать раздел *«Дисперсионный анализ»*, а в нем опцию — *«Двухфакторный»*. После этого в окне *«Анализ переменных»* нужно выбрать все 5 переменных (если в таблице нет лишних, то нажать на кнопку *«Все»*). После нажатия на кнопку *«Утвердить»* следует выбрать первый вариант из предложенных моделей *«1=неповторяемый»* и нажать на соответствующую кнопку.

Результаты анализа представляют собой стандартный набор показателей (см. табл. 21). Кроме того, STADIA в словесном виде формулирует результат проверки нулевых гипотез для каждого из двух факторов. В конце таблицы результатов выдаются значения параметров двухфакторной модели: общее среднее, отклонение от среднего, вызванное влиянием каждого из уровней фактора (эффекты 1—1 ... 1—5, 2—1 ... 2—5), и соответствующие им значения доверительных интервалов. В данной модели *не оценивается межфакторное взаимодействие*, что является ее серьезным ограничением.

Как следует из приведенных в таблице результатов, был обнаружен основной эффект фактора В — влияние возраста обучающихся. Основной эффект фактора А — методика обучения — оказался не значим, хотя величина F-отношения оказывается на грани статистической достоверности: $F = 2.99$, что чуть меньше $F_{кр.} = 3.01$ (значимость = 0.0502) и в принципе можно также говорить о влиянии и этого фактора.

Вариант ДА с повторными измерениями[1]. При обработке эксперимента с повторными измерениями каждому сочетанию уровней факторов соответствует несколько измерений независимой переменной. Исходные данные должны представлять собой псевдоматрицу, т.е. ее столбцы могут быть неодинаковой длины. Это означает, что на каждое сочетание факторов может приходиться различное количество измерений. Данные заносятся в электронную таблицу следующим образом: переменные (столбцы) соответствуют различным уровням обоих факторов и следуют в порядке изменения значения первого фактора: сначала все уровни первого фак-

[1] Варианту ДА с повторными измерениями (Repeated-Measures ANOVA) посвящена ниже отдельная глава. В системе Stadia данный алгоритм ДА рассматривается как один из вариантов простого двухфакторного ДА.

тора для первого уровня второго фактора, затем все уровни первого фактора для второго уровня второго фактора и т.д.

Таблица 21

Пример выдачи результатов двухфакторного дисперсионного анализа без повторных измерений статистической системой STADIA 6.1/prof

2-ФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ. Файл: educ55

Факторный план: неповторяемый

Источник	Сум.квадр	Ст.своб	Ср.квадр	Сила влияния
Факт.1	7.1E3	4	1.78E3	-0.282
Факт.2	1.34E4	4	3.36E3	0.333
Остат.	9.5E3	16	594	
Общая	3E4	24	1.25E3	

F(фактор1)=2.99, Значимость=0.0502, степ.своб = 4,16

Гипотеза 0: <Нет влияния фактора на отклик>

F(фактор2)=5.65, Значимость=0.0051, степ.своб = 4,16

Гипотеза 1: <Есть влияние фактора на отклик>

Параметры модели:

Среднее = 75.2, доверит.инт.=13

Эффект1-1 = 2.8, доверит.инт.=34.4

Эффект1-2 = -7.2, доверит.инт.=34.4

Эффект1-3 = 26.8, доверит.инт.=34.4

Эффект1-4 = -25.2, доверит.инт.=34.4

Эффект1-5 = 2.8, доверит.инт.=34.4

Эффект2-1 = -17.2, доверит.инт.=34.4

Эффект2-2 = -21.2, доверит.инт.=34.4

Эффект2-3 = -9.2, доверит.инт.=34.4

Эффект2-4 = 4.8, доверит.инт.=34.4

Эффект2-5 = 42.8, доверит.инт.=34.4

Рассмотрим *пример зависимости эффективности познавательной деятельности от курения и типа решаемой задачи* (см. *Хауэлл*, 1998). Двухфакторный эксперимент 3×3. В опытах участвовали испытуемые, отличавшиеся дозой принятого

никотина (фактор А, 3 уровня: A_1 — курили перед опытом, A_2 — курили за 3 часа до опыта, A_3 — не курят вообще). Испытуемые выполняли 3 вида когнитивных задач (фактор В, 3 уровня: B_1 — опознание целевого стимула, B_2 — задача на кратковременную память, B_3 — работа на автотренажере). В таблице 22 представлен результат ввода данных в электронную таблицу STADIA.

Таблица 22

Результаты выполнения заданий в эксперименте по исследованию эффекта курения на различные виды познавательной деятельности

	A_1B_1	A_1B_2	A_1B_3	A_2B_1	A_2B_2	A_2B_3	A_3B_1	A_3B_2	A_3B_3
1	9	27	15	12	48	7	8	34	3
2	8	34	2	7	29	0	8	65	2
3	12	19	2	14	34	6	9	55	0
4	10	20	14	4	6	0	1	33	0
5	7	56	5	8	18	12	9	42	6
6	10	35	0	11	63	17	7	54	2
7	9	23	16	16	9	1	16	21	0
8	11	37	14	17	54	11	19	44	6
9	8	4	9	5	28	4	1	61	4
10	10	30	17	6	71	4	1	38	1
11	8	4	15	9	60	3	22	75	0
12	10	42	9	6	54	5	12	61	0
13	8	34	3	6	51	16	18	51	6
14	11	19	15	7	25	5	8	32	2
15	10	49	13	16	49	11	10	47	3

Перед началом ДА следует проверить предположение о нормальности распределения данных и о равенстве дисперсий в отдельных группах. Для проведения этих тестов необходимо войти в главное меню статистических методов (**F9**) и выбрать раздел *«Параметрические тесты»*, в нем сначала пункт 2 — *«Гистограмма/нормальность»*, а затем пункт 4 — *«Тесты Стьюдента и Фишера»*. При проверке нормальности распределения данных нужно обработать каждую переменную (выбрав последовательно из списка и нажав на кнопку *«Утвердить»*). В ходе обработки STADIA оценит данные по трем статистическим критериям (Колмогорова—Смирнова, омега-квадрат и хи-квадрат) и выдаст после каждого теста словесное заключение об отличии проверяемого распреде-

ления от нормального, а также построит соответствующие гистограммы. Для оценки равенства дисперсий (пункт 4 по критерию Фишера в списке переменных) можно нажать на кнопку *«Все»*, и тогда STADIA проведет попарное сравнение всех переменных автоматически. После проведения теста с каждой парой переменных появится словесное заключение о равенстве/различии дисперсий выборок. В силу простоты выполнения данных статистических процедур, мы не будем их описывать и сошлемся лишь на конечный результат проведения этих проверок: распределения всех 9 переменных оказались не отличающимися от нормального, а их дисперсии равны. В том случае, если предположения о нормальности и равенстве дисперсий нарушаются (гистограммы асимметричны, а дисперсия явно увеличивается с ростом среднего), можно использовать один стандартный прием — сделать *математическое преобразование* с имеющимися данными. Обычно используют логарифмирование переменной или извлечения квадратного корня. Для выполнения преобразования переменных нужно выбрать в основном меню функций опцию *«Преобразования»* (**F8**), а в ней пункт 1 — *«Стандартная функция»*. Выбрав соответствующую кнопку *«1 = LOG»* или *«6 =* X^a*»*, можно осуществить нужное преобразование над той переменной, где стоит курсор. После этого приступить к выполнению ДА.

После проведенных проверок (или преобразований) приступаем непосредственно к проведению ДА. Как и в предыдущем примере, следует выбрать раздел *«Дисперсионный анализ»*, а в нем опцию — *«Двухфакторный»*. После этого в окне *«Анализ переменных»* нужно выбрать все 9 переменных (если в таблице нет лишних, то нажать на кнопку *«Все»*). После нажатия на кнопку *«Утвердить»* следует выбрать второй вариант ДА с повторными измерениями (*«Повторяемый»*) и модель с фиксированными эффектами (*«2 = фиксирован. эффекты»*). Кроме того (что очень важно!), в специальном окошке (*«Градаций фактора1 =* □) следует указать точное число уровней первого фактора, в нашем примере их 3.

Результаты обработки (см. табл. 23) показывают, что обнаружен основной эффект фактора А, т.е. эффективность дея-

тельности испытуемых зависела от принятия дозы никотина. Второй основной эффект оказался не значимым. Однако взаимодействие обоих факторов было высоко достоверным, значит курение влияет на разные познавательные процессы по-разному.

Таблица 23

Пример выдачи результатов двухфакторного дисперсионного анализа с повторными измерениями статистической системой STADIA 6.1/prof

2-ФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ. Файл: smoke

Факторный план: повторяемый, с фиксированными эффектами

Источник	Сум.квадр	Ст.своб	Ср.квадр	Сила влияния
Факт.1	2.87E4	2	1.43E4	0.0422
Факт.2	355	2	177	0.666
Межфакт.	2.73E3	4	682	-0.0615
Остат.	1.36E4	126	108	
Общая	4.53E4	134	338	

F(фактор1)=133, Значимость=0.0, степ.своб = 2,126

Гипотеза 1: <Есть влияние фактора на отклик>

F(фактор2)=1.64, Значимость=0.195, степ.своб = 2,126

Гипотеза 0: <Нет влияния фактора на отклик>

F(межфакт.)=6.33, Значимость=0.0002, степ.своб = 4,126

Гипотеза 1: <Есть влияние фактора на отклик>

Параметры модели:

Среднее = 569, доверит.инт.=61,4

Эффект1-1 = -552, доверит.инт.=159

Эффект1-1 = -550, доверит.инт.= 159

Эффект1-1 = -549, доверит.инт.= 159

Эффект2-1 = -559, доверит.инт.= 159

Эффект2-2 = -530, доверит.инт.= 159

Эффект2-3 = -562, доверит.инт.= 159

Статистическая система STADIA делает еще один полезный вид обработки данных для двухфакторного ДА с фик-

сированными или случайными эффектами. В том случае, если эффект межфакторного взаимодействия не обнаружен, то проводится дополнительный анализ по факторам А и В, но уже *без учета их взаимодействия.* Этот прием, как правило, дает более низкий уровень значимости нулевых гипотез, т.е. может повыситься вероятность обнаружения одного из факторных эффектов (как это было в предыдущем примере с основным эффектом фактора 2). Полученными результатами рекомендуется пользоваться, если уровень значимости гипотезы об отсутствии взаимодействия факторов достаточно высок (p значительно больше 0.05).

3.11. Работа с двухфакторным дисперсионным анализом в статистической системе SPSS 8.0.1

Данные заносятся в электронную таблицу *Редактора Данных* (SPSS Data Editor) в виде двух или нескольких столбцов (переменных) одинаковой длины: в первый заносятся значения исследуемой переменной, полученные по каждому из N наблюдений, во второй — значения уровней первого фактора (для a фиксированных уровней — это числа от 1 до k), соответствующие отдельным наблюдениям, в третий — значения уровней для второго фактора (для b фиксированных уровней — это числа от 1 до m). Таким образом, количество строк равно сумме *всех* наблюдений или сумме наблюдений, полученных по каждому из $k \times m$ уровней факторов. Если двухфакторный ДА выполняется последовательно для нескольких переменных, то электронная таблица, соответственно, может содержать большее количество столбцов.

Кроме того, в системе SPSS появляется еще одна дополнительная возможность обработки данных в процедуре многофакторного ДА. Это так называемая техника учета *ковариации* измеряемого признака с другим дополнительным признаком. Оставаясь в рамках нашего базового примера с исследованием зависимости кратковременной памяти от длины слов (см. разд. 3.4) и скорости их предъявления, предположим, что до эксперимента все испытуемые прошли стандартный тест на *объем кратковременной памяти*. Естественно предположить, что имеет место высокая корреляция (т.е.

линейная зависимость) между количеством правильно воспроизведенных букв испытуемыми каждой группы и результатами теста на кратковременную память. В этом случае (когда явная корреляция установлена!) в алгоритме расчетов ДА используется эта *дополнительная переменная-ковариата*. В ходе ДА метод использования переменной-ковариаты позволяет в ряде случаев увеличить величину основных эффектов и, тем самым, сделать их статистически значимыми. Таким образом, использование дополнительной информации повышает эффективность основной процедуры ДА.

Возьмем результаты одного инженерно-психологического эксперимента и обработаем имеющиеся данные в системе SPSS 8.0.1[1]. Воспользуемся данными из задания № 4 предыдущей главы (см. с.41—42). Данные (см. табл. 24) представляют собой количество ошибок, допущенных тремя группами испытуемых (фактор 1: 3 уровня) в эксперименте по слежению за движущимся объектом. Кроме того, введем еще один контролируемый фактор — пол испытуемого (2 уровня). До начала выполнения вычислений проделаем небольшой визуальный анализ полученных результатов. Представленный на рис. 8 своеобразный «параллелограмм ошибок» показывает, что фактор 1 (группа испытуемых), по-видимому, влиял на количество допущенных ошибок, поскольку виден рост количества ошибок от первой группы к третьей. Кроме того, можно предположить, что фактор 2 (пол испытуемых) также оказывал влияние на успешность деятельности (по крайней мере, это видно во второй группе). По поводу наличия межфакторного взаимодействия что-то определенное сказать трудно, но заранее отвергать его, по-видимому, нет оснований, поскольку в двух группах (новички и студенты) мужчины и женщины явно отличаются по количеству сделанных ошибок. Посмотрим, насколько наш пилотажный визуальный анализ будет соответствовать результатам ДА.

---

[1] По сравнению с более ранними версиями SPSS, в версиях 8.0 и выше в процедурах ДА были сделаны значительные изменения. Тем не менее, мы надеемся, что читатель, работавший с более ранними версиями данной системы, сможет без особого труда понять смысл сделанных изменений.

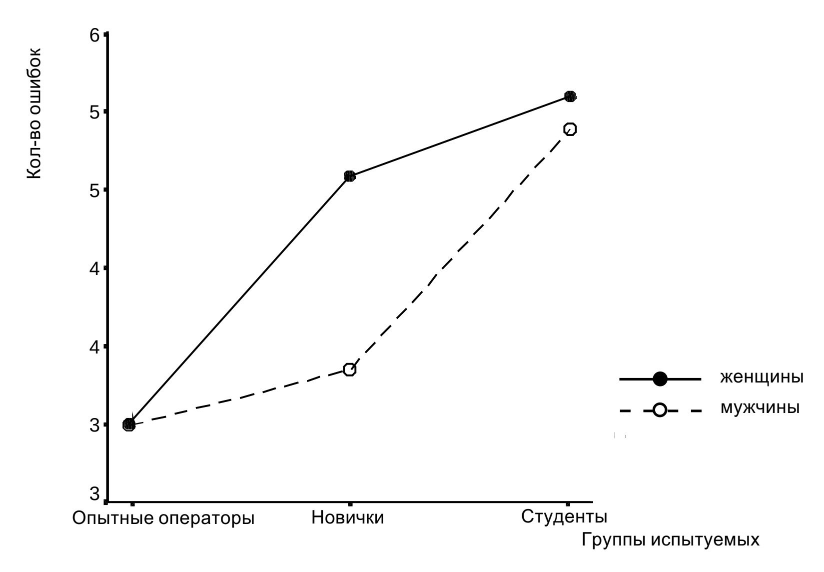


Рис. 8. Результаты двухфакторного эксперимента по слежению за движущимся объектом:
ось X — группа испытуемых; ось Y — количество допущенных ошибок

После ввода данных в меню статистических методов *Статистика* (Statistics) следует выбрать группу методов ОЛМ — *Общая линейная модель* (General Linear Model), а ней вариант *Обобщенная факторная...* (General Factorial). В появившемся окошке слева будет представлен список всех переменных (столбцов), включенных в электронную таблицу. В нашем случае их 4: *№ испытуемого, Ошибки, Группа и Пол.* Из них нужно выбрать *одну* анализируемую переменную (*Ошибки*) и перенести ее в окно *Зависимая переменная* (Dependent Variable).

Затем следует указать, в каких переменных заданы значения уровней исследуемых факторов, и перенести их в окно *Фиксированные факторы* (Fixed Factors) или *Случайные факторы* (Random Factors). В нашем примере имеется 2 фиксированных фактора — *Группа* и *Пол.*

В данном примере мы не используем переменную-ковариату, поэтому в окно *Ковариаты* (Covariates) не заносится никакая переменная. В следующем окне *Взвешенный МНК* (WLS Weight) можно задать переменную, значения которой придают различные веса результатам измерений, составляю-

Таблица 24

Данные эксперимента по слежению за движущимся объектом, внесенные в электронную таблицу SPSS

№	Ошибки	Группа	Пол
1	3.13	1.00	1.00
2	3.64	1.00	1.00
3	2.59	1.00	1.00
4	4.22	1.00	1.00
5	1.36	1.00	1.00
6	3.25	1.00	2.00
7	3.14	1.00	2.00
8	1.97	1.00	2.00
9	3.16	1.00	2.00
10	3.47	1.00	2.00
11	1.39	2.00	1.00
12	3.87	2.00	1.00
13	4.37	2.00	1.00
14	3.79	2.00	1.00
15	3.33	2.00	1.00
16	5.38	2.00	2.00
17	4.07	2.00	2.00
18	5.39	2.00	2.00
19	3.37	2.00	2.00
20	4.74	2.00	2.00
21	5.60	3.00	1.00
22	3.02	3.00	1.00
23	6.18	3.00	1.00
24	5.22	3.00	1.00
25	4.15	3.00	1.00
26	5.47	3.00	2.00
27	6.88	3.00	2.00
28	6.40	3.00	2.00
29	2.07	3.00	2.00
30	4.68	3.00	2.00

щим зависимую переменную. Эта возможность заключается в том, что числовые значения зависимой переменной могут быть трансформированы посредством *процедуры взвешивания* по методу наименьших квадратов: WLS Weight (Weighted Least-Squares). Для этого в данных можно указать отдельную переменную, с помощью которой результаты проведенных измерений получат различный вес. Подобная трансформация, как правило, выполняется при необходимости компенсировать различия в точности измерений используемых в исследовании переменных. Тем, кто не знаком с использованием данного приема, мы не рекомендуем его применение. Поскольку в нашем эксперименте анализируется толь-

ко одна зависимая переменная, то мы не используем данную возможность.

Далее пользователь может определить варианты используемой модели ДА, для чего следует нажать кнопку *Модель* (Model). *Полная факторная* модель (Full factorial) включает в себя оценки всех главных факторных эффектов, эффекты ковариат, парные взаимодействия всех факторов и не включает оценку взаимодействия между ковариатами. Чтобы задать другую необходимую пользователю конфигурацию модели (т.е. ограничить полную факторную модель), следует при выборе модели указать вариант *Настраиваемая* (Custom) и определить все необходимые компоненты модели. При обработке более чем двухфакторного эксперимента исследователь может *ограничить уровень межфакторного взаимодействия*, т.е. при расчетах отсечь взаимодействие факторов высшего порядка. В случае введенного ограничения любой не вычисляемый эффект взаимодействия будет объединен с остаточной суммой квадратов. В случае задания отсутствия взаимодействий будут рассчитываться *только основные эффекты* по каждому из факторов. Поскольку в текущем анализе нас интересуют не только основные эффекты, но и взаимодействие факторов, мы, без сомнения, выбираем полную факторную модель.

Кроме того, пользователь имеет возможность выбрать один из трех методов разложения суммы квадратов и оценки остаточной суммы квадратов (основная процедура ДА) — окно *Сумма квадратов* (Sum of squares). Для сбалансированных и несбалансированных моделей[1], как правило, используется метод типа III, который также в литературе по ДА называют *Уникальным* (Unique). Поскольку этот метод инвариантен относительно числа наблюдений по отдельным уровням факторов, он чаще всего применяется на практике. При использовании ковариаты могут быть выбраны также три других

---

[1] Напомним, что в так называемой сбалансированной модели ДА для каждого уровня фактора число наблюдений равно, для несбалансированной модели — различно.

метода — тип I (*Иерархический*, Hierarhical), тип II (*Экспериментальный*, Experimental) или тип IV. Для осмысленного выбора одного из них следует обратиться к специальной литературе, поскольку даже обширное описание процедур ДА, предлагаемое фирмой SPSS, не дает пользователю четких рецептов, и это не входит также в нашу задачу. Не очень опытным пользователям мы советуем выбрать тип III, заданный по умолчанию. Тем же, кто в достаточной степени освоил ДА и интересуется его нюансами, использование различных методов оценки суммы квадратов позволит оценить специфическое влияние переменных-ковариат на выраженность факторных эффектов и факторных взаимодействий, а также с большей адекватностью учесть особенности собственного экспериментального плана. Для нашего примера мы выбираем метод типа III; количество межфакторных взаимодействий ограничивать не стоит, поскольку у нас и так имеется лишь два фактора.

Обычно в модель ДА включают свободный член, поэтому в окошке *Включить в модель свободный член* (Incude intercep in model) поставим «галочку».

Нажав на кнопку *Контрасты* (Contrasts), пользователь может задать 5 различных видов оценки априорного контраста. Поскольку у нас нет конкретных априорных гипотез, то эту опцию мы не используем. Если хотя бы один из факторов включает больше, чем 2 уровня, то для него (для них) можно заказать оценку множественного сравнения средних, для чего следует нажать на кнопку *Апостериори* (Post Hoc), указать необходимый фактор и выбрать нужные статистические тесты. Мы указываем фактор *Группа* и выбираем один из тестов — тест Шеффе.

В качестве дополнительных статистических показателей закажем в меню *Параметры* (Options) расчет описательных статистик, оценку силы факторных эффектов и проверку однородности дисперсий.

Рассмотрим результаты ДА, представленные в таблице 25. Первая часть результатов анализа представляет собой краткий

обзор введенных данных (*Межобъектные факторы*[1], Between-Subject Factors). Далее — тесты описательных статистик (*Дескриптивные статистики*, Descriptive Statistics). Еще ниже — результаты проверки равенства дисперсий (Levene‘s Test of Equality of Variances) — $F = 1.143$, $p = 0.365$ (очевидно, что групповые дисперсии достоверно не отличаются друг от друга). Вторая часть — непосредственно результаты ДА или оценка значимости факторных эффектов (*Проверки межобъектных эффектов*, Tests of Between-Subjects Effects). Значимым оказался основной эффект первого фактора — «Группа испытуемых». Второй фактор («Пол испытуемых»), и межфакторное взаимодействие оказались незначимыми. Интересно, что простой графический анализ результатов эксперимента, проведенный нами до выполнения ДА, явно не соответствует строгим результатам ДА: не подтвердился намек на основной эффект второго фактора, да и межфакторное взаимодействие оказалось незначимым.

Приведенные ниже результаты тестов парных сравнений средних по группам испытуемых показывают (см. те строки, которые помечены «звездочкой»), что факторный эффект обусловлен различиями между группой операторов-профессионалов и группой студентов.

Попробуем исключить из расчетов эффект межфакторного взаимодействия (Maximum Interactions = None) и еще раз проведем ДА. Самая нижняя часть таблицы 25 показывает, что немного увеличился основной эффект первого и второго факторов ($F_{GROUP} = 6.79$ против 6.64 и $F_{SEX} = 1.33$ против 1.30). Таким образом, исключение межфакторного взаимодействия нам практически ничего не дало.

---

[1] В настоящем учебном пособии мы используем, на наш взгляд, более удачное название *межгрупповые факторы*.

Таблица 25

Пример выдачи результатов двухфакторного дисперсионного анализа с повторными измерениями статистической системой SPSS 8.0.1

Однофакторный дисперсионный анализ

Межгрупповые факторы

		Метка значения	N
Группа испытуемых	1	Операторы-профессионалы	10
	2	Операторы-новички	10
	3	Студенты	10
Пол испытуемого	1	женщины	15
	2	мужчины	15

Дескриптивные статистики

Зависимая переменная: Количество ошибок

Группа испытуемых	Пол испытуемого	Среднее	Стд. отклонение	N
Операторы-профессионалы	женщины	2.9880	1.0922	5
	мужчины	2.9980	.5894	5
	Итог	2.9930	.8274	10
Операторы-новички	женщины	3.3500	1.1561	5
	мужчины	4.5900	.8726	5
	Итог	3.9700	1.1660	10
Студенты	женщины	4.8340	1.2556	5
	мужчины	5.1000	1.8942	5
	Итог	4.9670	1.5215	10
Итог	женщины	3.7240	1.3626	15
	мужчины	4.2293	1.4835	15
	Итог	3.9767	1.4229	30

Критерий Ливиня проверки равенства дисперсий.

Зависимая переменная: Количество ошибок

F	ст.св.1	ст.св.2	Знач.
1.143	5	24	.365

Оценка факторных эффектов

Зависимая переменная: Количество ошибок

Источник	Сумма квадратов типа III	ст.св.	Средний квадрат	F	Знач.	Эта в квадрате
Скорректированная модель	23.505[a]	5	4.701	3.204	.023	.400
Intercept	474.416	1	474.416	323.358	.000	.931
ГРУППА	19.484	2	9.742	6.640	.005	.356
ПОЛ	1.915	1	1.915	1.305	.265	.052
ГРУППА * ПОЛ	2.106	2	1.053	.718	.498	.056
Ошибка	35.212	24	1.467			
Итог	533.133	30				
Скорректированный итог	58.717	29				

a. R квадрат = .400 (Скорректированный R квадрат = .275)

Таблица 25 (продолжение)

Апостериорные критерии

Множественные сравнения

Зависимая переменная: Количество ошибок
Шеффе

(I) Группа испытуемых	(J) Группа испытуемых	(I-J)-я разность средних	Стд. ошибка	Знач.	95% доверительный интервал Нижняя граница	95% доверительный интервал Верхняя граница
Операторы-профессионалы	Операторы-новички	-.9770	.542	.218	-2.3901	.4361
	Студенты	-1.9740*	.542	.005	-3.3871	-.5609
Операторы-новички	Операторы-профессионалы	.9770	.542	.218	-.4361	2.3901
	Студенты	-.9970	.542	.205	-2.4101	.4161
Студенты	Операторы-профессионалы	1.9740*	.542	.005	.5609	3.3871
	Операторы-новички	.9970	.542	.205	-.4161	2.4101

Основано на наблюденных средних.
*. Разность средних значима на уровне .05.

Проверки факторных эффектов без учета межфакторного взаимодействия

Зависимая переменная: Количество ошибок

Источник	Сумма квадратов типа III	ст.св.	Средний квадрат	F	Знач.
Скорректированная модель	21.399	3	7.133	4.970	.007
Свободный член	474.416	1	474.416	330.536	.000
ГРУППА	19.484	2	9.742	6.787	.004
ПОЛ	1.915	1	1.915	1.334	.259
Ошибка	37.318	26	1.435		
Итог	533.133	30			
Корректированный итог	58.717	29			

Пояснения к основным терминам в таблице:

1. Кроме основных эффектов каждого фактора (ГРУППА, ПОЛ), SPSS также оценивает их совместный эффект (ГРУППА×ПОЛ).
2. Термин «Скорректированная Модель (Corrected Model)» соответствует использованному выше термину «общая факторная» или межгрупповая часть «общей дисперсии». Ошибка (Residual) обозначает «остаточную» часть общей дисперсии, а Итог (Total) — «общую» дисперсию.

Задания

1. На рисунках 9—14 показаны соотношения средних значений зависимой переменной, полученные в шести разных двухфакторных экспериментах 3×2. Подумайте, в каких случаях можно предполагать наличие межфакторного взаимодействия, а в каких нет.

2. Заполните до конца таблицу ДА и выполните по *F*-критерию проверку нулевых гипотез для факторов А и В и их взаимодействия.

Источник вариации		Степени свободы, d_f	$\sum d^2$	$\hat{s}^2$	*F*
Главные эффекты	По фактору А	4	64.26	?	?
	По фактору В	5	46.85	?	?
	Взаимодействие А× В	?	?	?	?
	Остаточная	120	1136.53	?	—
	Общая	149	?	?	—

3. Обработайте в одной из статистических систем данные следующего эксперимента 5×2 и дайте интерпретацию результатов анализа. Две группы испытуемых — экстраверты и интраверты, отобранные предварительно по опроснику Г.Айзенка (фактор А: 2 уровня), участвовали в опытах по исследованию эффективности запоминания словесного материала в пяти различных стрессогенных условиях (фактор В: 5 уровней). В таблице 26 представлено количество правильно воспроизведенных слов для всех 10 сочетаний факторов.

Таблица 26

Данные для обработки к заданию 2

Стресс 1		Стресс 2		Стресс 3		Стресс 4		Стресс 5	
Интр.	Экстр.	Интр.	Экстр.	Интр.	Экстр.	Интр.	Экстр.	Интр.	Экстр.
9	8	7	10	11	14	12	20	10	21
8	6	9	7	13	11	11	16	19	19
6	4	6	8	8	18	16	16	14	17
8	6	6	10	6	14	11	15	5	15
10	7	6	44	14	13	9	18	10	22
4	6	11	7	11	22	23	16	11	16
6	5	6	10	13	17	12	20	14	22
5	7	3	6	13	16	10	22	15	22
7	9	8	7	10	12	19	14	11	18
7	7	7	7	11	11	11	19	11	21

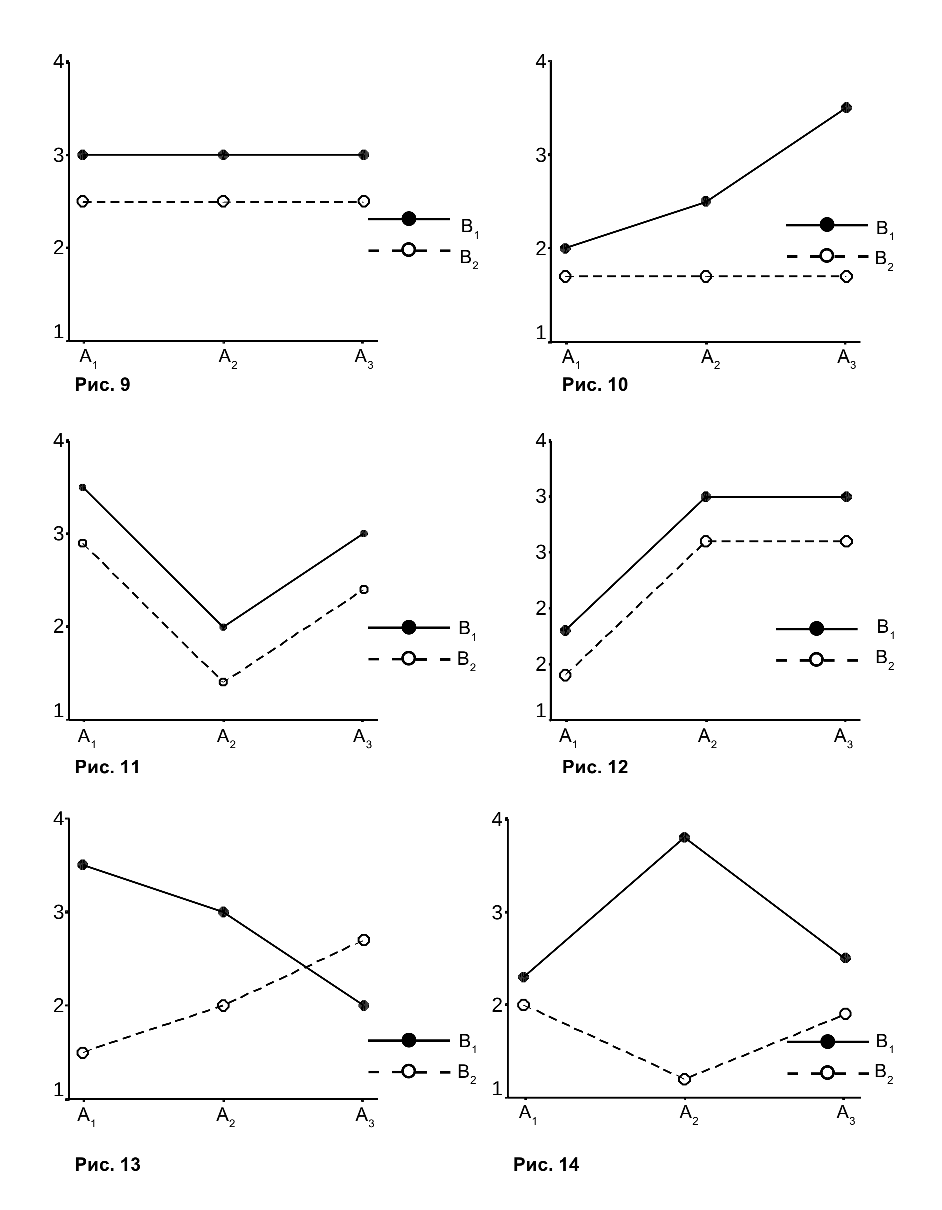


Рис. 9—14. Различные варианты зависимости измеряемой переменной от факторов А и В:

ось X — уровни фактора А; ось Y — величина зависимой переменной в условных единицах; сплошная линия — зависимость при первом уровне фактора В; пунктир — при втором уровне фактора В

4. Обработайте в одной из статистических систем данные следующего трехфакторного эксперимента 2×3 (см. табл. 27) и дайте их интерпретацию. На компьютерном автотренажере изучались навыки вождения у двух групп испытуемых — новичков и опытных водителей (фактор А: 2 уровня); испытуемые должны были проехать трассу трех классов сложности (фактор В: 3 уровня) в двух различных условиях вождения автомобиля — днем и ночью (фактор С: 2 уровня). В таблице 27 представлено количество ошибок, допущенных водителями в каждом из условий. Предварительно постройте и проанализируйте графики, сформулируйте свои гипотезы о выраженности основного эффекта каждого из факторов и их взаимодействиях.

Таблица 27

Данные для обработки к заданию 4

Факторы А, В и С	День			Ночь		
	B_1	B_2	B_3	B_1	B_2	B_3
Новички	4	23	16	21	25	32
	18	15	27	14	33	42
	8	21	23	19	30	46
	10	13	14	26	20	40
Опытные	6	2	20	11	23	17
	4	6	15	7	14	16
	13	8	8	6	13	25
	7	12	17	16	12	12

5. Проведите ДА данных задания 4, последовательно исключив из анализа тройное, а затем двойное межфакторное взаимодействие. Как это повлияло на величину *F*-отношения для каждого из факторов?

6. Оцените силу факторных эффектов в примере упражнения 2.

7. По результатам проведенного ДА в упражнении 3 оцените степень адекватности линейной модели ДА полученным данным.

8. Найдите в литературе пример двухфакторного эксперимента[1], где автор работы не применял ДА. В том случае, если автором приводятся лишь средние значения зависимой переменной, создайте данные сами (так, чтобы они соответствовали приведенным средним) и обработайте их в одной из статистических систем. Проанализируйте полученные результаты и подумайте, согласны ли вы с основными выводами автора.

[1] Для этой цели стоит обратиться к любому номеру «Психологического журнала» за 70—90-е годы и найти там экспериментальную статью.

Глава 4

Дисперсионный анализ данных с повторными измерениями

4.1. Особенности ДА с повторными измерениями

В предыдущем разделе мы рассматривали примеры, когда в разных экспериментальных группах (клетках таблицы для ДА) были результаты измерений зависимой переменной, полученные от *разных* испытуемых. Строго говоря, мы имели дело с *независимыми или несвязанными* выборками. Однако нередко бывает, что используется экспериментальный план, при котором в различных сочетаниях уровней факторов участвовали одни и те же испытуемые. Например, в эксперименте по исследованию эффективности совместной деятельности варьировали уровень мотивации (фактор А: 3 уровня) и оценивали влияние проводимого тренинга общения — до и после него (фактор В: 2 уровня). В подобных экспериментах, как правило, используют данные одной и той же группы испытуемых. Но в таком случае трудно предположить, что повторные данные независимы и не коррелируют друг с другом, поскольку самый успешный испытуемый покажет более высокие результаты при различных экспериментальных условиях, а неуспешный будет в среднем так же ниже остальных. Это означает, что мы нарушаем очень важное допущение ДА о независимости выборок.

Для компенсации этого недостатка в ДА используют метод оценки и исключения влияния зависимости выборок, вызванного повторным использованием одних и тех же испытуемых в разных экспериментальных условиях. В англоязычной литературе этот вариант многофакторного ДА получил название *Repeated-Measures Design, т.е. эксперименталь-*

ный план с повторными измерениями. Вариант ДА с повторными измерениями имеет большие преимущества, поскольку использование одних и тех же испытуемых во всех группах позволяет уменьшить общую дисперсию данных за счет *исключения компоненты межиндивидуальных различий* из остаточной дисперсии, что, в свою очередь, позволяет получать более высокие оценки *F*-отношения и с большей легкостью отвергать нулевые гипотезы.

В качестве иллюстрации эффекта взаимодействия выборок рассмотрим специально подготовленный пример с нарочито преувеличенными данными. Пусть проведен однофакторный эксперимент, в котором измерялось число правильных ответов и варьировались 3 уровня фактора А (см. табл. 28).

Таблица 28

Данные однофакторного эксперимента с использованием одной и той же группы испытуемых

Испытуемый	Экспериментальные условия (уровни фактора)			Всего по испытуемому
	A_1	A_2	A_3	
1	21	43	71	136
2	106	120	128	356
3	232	292	305	832
4	305	312	338	956
Всего по условию	664	767	842	2280

Как видно, суммы по столбцам (по условиям) отличаются незначительно, и эти вариации скорее всего носят случайный характер. Однако суммы по строкам, характеризующие работу каждого из 4 испытуемых, отличаются значительно и характеризуют индивидуальный уровень исполнения тестового задания. Испытуемый 1 имеет меньше всего правильных ответов, и это закономерно повторяется во всех экспериментальных условиях. Испытуемые 3 и 4 значительно более успешны, что также проявилось на всех уровнях. Таким образом, вариабельность внутри условий эксперимента

в основном является *следствием индивидуальных различий* между испытуемыми, а не следствием различий трех экспериментальных условий. Таким образом, если бы мы смогли исключить эти индивидуальные различия, мы смогли бы более точно оценить ошибку дисперсионной модели и уменьшить ее. Кроме того, индивидуальные различия, создающие такую высокую корреляцию между экспериментальными условиями, могут быть выделены в качестве особого фактора, характеризующего испытуемых.

На практике в вычислительных алгоритмах ДА с повторными измерениями сумма квадратов, соответствующая межиндивидуальной вариации, вычисляется и вычитается из общей суммы квадратов до начала вычислений остальных сумм квадратов. Эта сумма квадратов получила название межиндивидуальной суммы квадратов — $\sum d_{BS}^2$. Индекс *BS* означает between subjects, т.е. «между испытуемыми». Таким образом, в модели ДА с повторными измерениями мы обозначили еще один компонент общей дисперсии, обусловленный влиянием межиндивидуальных вариаций испытуемых.

4.2. Структурная модель ДА с повторными измерениями

Линейная модель ДА, описанная выше (см. с.9), дополняется влиянием индивидуальных различий испытуемых:

$$X_{ij} = \mu_{общ.} + \tau_j + p_i + \tau p_{ji} + \varepsilon_{ij} \quad (38)$$

где $\mu_{общ.}$ — общее среднее или свободный член уравнения, τ_j — дополнительный компонент, соответствующий влиянию *j*-го фактора (т.е. насколько среднее *j*-го фактора отличается от среднего генеральной совокупности), p_i — компонент, связанный с влиянием индивидуальности *i*-го испытуемого, τp_{ji} — дополнительный эффект взаимодействия этих двух компонентов, ε_{ij} — ошибка линейной модели, связанная с *j*-м фактором и *i*-м испытуемым.

4.3. Статистические критерии и предположения, используемые в ДА с повторными измерениями

Как уже было описано выше, в моделях одномерного ДА, где исследуется только влияние межгрупповых факторов, для оценки факторного эффекта используется обычный *F*-критерий или, как его еще называют, одномерный *F*-тест. В ДА с повторными измерениями используются два подхода — *одномерный и многомерный.*

Одномерный подход (его еще называют моделью расщепленных делянок[1] или смешанной моделью) рассматривает зависимые переменные как повторные измерения, проведенные на различных уровнях внутригруппового фактора. Предполагается, что каждое измерение, полученное на отдельном испытуемом (объекте измерения), должно быть выборкой из многомерного нормального распределения. Кроме того, следует отметить необходимость выполнения еще одного математического допущения, чтобы быть уверенным, что любое эмпирическое *F*-отношение соответствует стандартному (табличному) распределению *F*-отношения. Это допущение о так называемой составной симметрии (compound symmetry) ковариационной матрицы, образованной переменными, соответствующими уровням межгрупповых факторов. Таблица 29 позволит нам точнее понять, о чем идет речь. Если мы имеем один фактор с тремя уровнями (как в предыдущем примере), расчет коэффициентов ковариации между столбцами матрицы данных (см. табл. 29: всего 3×3 = 9 элементов) дает нам ковариационную матрицу. Выполнение этого допущения означает, что элементы матрицы, лежащие на основной диагонали, должны быть равны, и остальные элементы (лежащие вне основной диагонали) также должны быть одинаковы. Отметим также, что равенство *диагональных элементов* корреляционной матрицы ($cov_{11} = cov_{22} = cov_{33}$) означает выполнение допущения о *гомогенности дисперсий* выборок.

---

[1] Исторически так сложилось, что изначально ДА применялся для оценки различных экспериментальных эффектов в сельском хозяйстве. Отсюда и термин «делянка», т.е. та часть экспериментального поля, где оценивается определенный факторный эффект. В психологии «делянке», как правило, соответствует отдельная группа испытуемых, дифференцируемая от других по определенному признаку — межгрупповому фактору.

Таблица 29

Ковариационная матрица для данных однофакторного эксперимента со связанными выборками

	A_1	A_2	A_3
A_1	cov_{11}	cov_{12}	cov_{13}
A_2	cov_{21}	cov_{22}	cov_{23}
A_3	cov_{31}	cov_{32}	cov_{33}

Это допущение о составной симметрии ковариационной матрицы является частью более общего допущения о форме дисперсионно-ковариационных матриц, образованных межгрупповыми факторами и зависимыми переменными (внутригрупповыми факторами). Предполагается, что дисперсионно-ковариационные матрицы, полученные по межгрупповым факторам — одинаковы, а по зависимым переменным — это единичная матрица. Только при выполнении таких допущений использование *F*-критерия становится надежным средством статистической оценки факторных эффектов и межфакторных взаимодействий в ДА с повторными измерениями. Для проверки последнего предположения обычно используется тест сферичности Моучли (Mauchly). Тем не менее, не стоит переоценивать значение использования данного теста, поскольку для малых выборок он не обладает большой мощностью, а для больших выборок даже очень небольшие отклонения могут оказаться статистически значимыми. Когда допущение о сферичности оказывается нарушенным, в процедуре ДА с повторными измерениями может вводиться специальная поправка, уменьшающая число степеней свободы числителя и знаменателя в *F*-отношении, и, следовательно, значимость *F*-критерия оценивается со скорректированными степенями свободы. В ряде статистических систем такая поправка обозначается как *эпсилон-коррекция* (epsilon ajustment).

Многомерный подход также рассматривает зависимые переменные как повторные измерения, проведенные на различных уровнях внутригруппового фактора. Для статистической оценки факторного эффекта в этом подходе использует-

ся не *F*-критерий, а так называемые многомерные тесты (подробнее об этом см. в следующей главе, посвященной многомерному ДА). В этом подходе также предполагается, что измерения, полученные по каждому испытуемому, должны быть выборкой из многомерного нормального распределения. Кроме того, принимается допущение об идентичности дисперсионно-ковариационных матриц, образованных межгрупповыми факторами. Для проверки предположения об идентичности дисперсионно-ковариационных матриц в большинстве статистических систем обычно используется М-тест Бокса (Box's M-test).

Поскольку в ДА с повторными измерениями для оценки факторного эффекта используются как одномерный, так и многомерные критерии, правомерен вопрос: какой подход более предпочтителен? В том случае, если оба подхода дают сходные результаты, вопрос о выборе не так уж важен. Однако, в случае, когда один из подходов оценивает факторный эффект как значимый, а другой — нет, поставленный выше вопрос весьма важен. Если для обоих подходов указанные выше допущения удовлетворяются, то в целом одномерный подход обладает большей мощностью по сравнению с многомерным. Это означает, что он с большей вероятностью обнаружит оцениваемый факторный эффект или взаимодействие факторов. Это преимущество особенно заметно на малых выборках. Вместе с тем, заметим, что многомерный подход в целом накладывает меньше ограничений на анализируемые данные.

4.4. Пример обработки однофакторного эксперимента с повторными измерениями

Покажем, как делаются все вычисления по данным эксперимента из области психосоматики (см. *Хауэлл*, 1998). На группе из 9 испытуемых, страдающих мигренью, изучалось влияние тренинга релаксации на частоту проявления головных болей. Испытуемых просили регистрировать количество часов мигренеподобных приступов в течение 5 недель: 2 недели до тренинга (фоновые замеры) и 3 недели в ходе тренинга (воздействие).

Таблица 30

Данные о частоте головных болей

Испытуемый	Фон		Тренинг			Всего по испытуемым
	1 неделя	2 неделя	3 неделя	4 неделя	5 неделя	
1	21	22	8	6	6	63
2	20	19	10	4	4	57
3	17	15	5	4	5	46
4	25	30	13	12	17	97
5	30	27	13	8	6	84
6	19	27	8	7	4	65
7	26	16	5	2	5	54
8	17	18	8	1	5	49
9	26	24	14	8	9	81
Всего по неделям	*201*	*198*	*84*	*52*	*61*	
В целом по эксперименту	N_{total} = 45,		Σ_{total} = 596,		s^2_{total} = 71.96	

Сначала рассчитаем *общую* по всем испытуемым сумму квадратов:

$$\sum d^2_{total} = (N_{total} - 1) \cdot s^2_{total} = 44 \cdot 71.96 = 3166.24. \quad (39)$$

Затем *межиндивидуальную* сумму квадратов:

$$\sum d^2_{IWG} = \sum_{j=1}^{n} \frac{(\sum X_i)^2}{k} - \frac{(\sum_{i=1}^{N} X_i)^2}{N} =$$

$$= \frac{63^2 + 57^2 + \ldots + 81^2}{5} - \frac{(596)^2}{45} = \quad (40)$$

$$= 8380.40 - 788.48 = 486.71,$$

где *n* — количество испытуемых в каждой группе, *k* — число недель наблюдения (групп), *j* — знак суммирования средних по испытуемым, а *i* — индекс суммирования значений всех наблюдений в отдельности. Индекс *IWG* означает individual differences within group, или межиндивидуальные различия внутри одной группы.

Далее вычислим *межгрупповую (или факторную)* сумму квадратов:

$$\sum d_{BG}^2 = \sum_{m=1}^{k} \frac{(\sum X_i)^2}{n} - \frac{(\sum_{i=1}^{N} X_i)^2}{N} =$$
$$= \frac{201^2 + 198^2 + ... + 61^2}{9} - \frac{(596)^2}{45} = \qquad (41)$$
$$= 10342{,}89 - 788.48 = 2449.20,$$

где n — число испытуемых в группе, m — знак суммирования средних по k группам.

И наконец, *внутригрупповую (или остаточную)* сумму квадратов вычисляем как разность между общей факторной, межиндивидуальной и межгрупповой суммами квадратов:

$$\sum d_{WG}^2 = \sum d_{total}^2 - \sum d_{IWG}^2 - \sum d_{BG}^2 =$$
$$= 3166.24 - 486.71 - 2449.20 = 230.33 \, . \qquad (42)$$

Таким образом, при оценке остаточной суммы квадратов мы *исключили влияние* межиндивидуальной вариативности и тем самым ее уменьшили. Результаты проведенного ДА показывают статистическую значимость контролируемого фактора — *неделя* эксперимента, что вполне соответствует простому сравнению суммарных значений времени головных болей, данных в последней строке таблицы 30. Поэтому у нас есть все основания думать, что проведенный тренинг релаксации был успешным. Дополнительный статистический анализ парных сравнений средних по группам будет дан ниже при проведении ДА в системе SPSS.

Сведем все дальнейшие вычисления в таблицу результатов ДА (см. табл. 31).

Эффективность использования модели ДА с повторными измерениями можно продемонстрировать, проведя однофакторный ДА при условии, что в исследовании участвовали 3 *различные группы* испытуемых. В этом случае величина F-отношения для исследуемого фактора будет равна 34.15.

Таблица 31

Результаты однофакторного ДА с повторными измерениями

Источник вариации	Степени свободы	$\sum d^2$	$\hat{s}^2$	F	P
Межиндивидуальная	8	486.71			
Факторная (межгрупповая)	4	2449.20	612.30	**85.04**	**<0.01**
Остаточная (внутригрупповая)	32	230.40	7.20		
Общая	44	3166.24			

Сравнив эту величину с аналогичной из таблицы 31, мы увидим, что она *более чем вдвое меньше*. Таким образом, мы получили очень большой выигрыш при оценке факторного эффекта.

Стоит обратить внимание на отсутствие в таблице результатов оценки взаимодействия межиндивидуального и факторного источников вариации, как следовало бы из описанной выше структурной модели ДА с повторными измерениями. Фактически в том случае, когда имеется лишь одно измерение в клетке дисперсионной таблицы (см. табл. 30, где на пересечении строки «испытуемый» со столбцом «неделя» приходится всего одно число), сумма квадратов взаимодействия совпадает с суммой квадратов остатков. В данном варианте экспериментального плана эти два источника вариации разделить невозможно.

Чтобы читатель лучше ориентировался в зарубежных статистических системах и их описаниях, определим, что фактор, контролируемый на одной и той же выборке испытуемых, в англоязычной литературе принято называть *Within-Subjects Factor* или *Within-Subjects Variable*. Мы называем его ***внутригрупповым фактором,*** подчеркивая тем самым, что его спецификой является ***интраиндивидуальная вариабельность*** одной и той же группы испытуемых (subjects), наблюдаемых повторно, т.е. при различных уровнях внутригруппового фактора. При оценке влияния таких факторов учитываются индивидуальные различия. В отличие от него фактор, заданный на выборках разных испытуемых, называют *Between-Subjects*

Factor или *Between-Subjects Variable*. Мы называем его *межгрупповым фактором*, подчеркивая, что его спецификой является *интериндивидуальная вариабельность*, наблюдаемая в группах, состоящих из различных испытуемых, в соответствии с заданными уровнями фактора. Эти факторы разделяют выборку испытуемых на отдельные группы. При оценке влияния межгрупповых факторов индивидуальные различия не учитываются.

Если мы в последнем примере для наглядной проверки сделанного предположения о симметричности ковариационной матрицы визуально оценим ее, то увидим, что это предположение резонно принять, поскольку разброс сравниваемых коэффициентов ковариации не столь уж велик: для диагональных элементов он составил от 11.5 до 28.5, а для недиагональных — 7.33—16.38 (см. табл. 32).

Таблица 32

Ковариационная матрица, рассчитанная по экспериментальным данным

	Неделя 1	Неделя 2	Неделя 3	Неделя 4	Неделя 5
Неделя 1	21.00	11.75	9.25	7.83	7.33
Неделя 2	11.75	28.50	13.38	16.38	13.38
Неделя 3	9.25	13.75	11.50	8.58	8.21
Неделя 4	7.83	16.38	8.58	11.69	10.82
Неделя 5	7.33	13.38	8.21	10.82	16.95

Тест Моучли в этом примере не обнаружил нарушения сферичности.

4.5. Двухфакторный ДА с внутригрупповыми и межгрупповыми факторами

Рассмотренный выше пример представляет собой простой однофакторный эксперимент. В реальной практике экспериментальных исследований чаще встречаются более сложные экспериментальные планы, когда контролируются и внутригрупповые и межгрупповые факторы. Такие экспериментальные планы иногда называют *смешанными*.

Пример 1: один фактор внутригрупповой, другой — межгрупповой[1].

Канадский исследователь Д.А.Кинг (1986) проводил эксперимент по изучению привыкания крыс к воздействию наркотиков. Использовались три *различные группы* по 8 животных (фактор А — «группа»: 3 уровня) — контрольная и две экспериментальных. Поскольку использованный наркотик оказывает быстрое действие, особенности двигательного поведения животных в каждой группе оценивались в течение часа после его приема по 6 пятиминутным интервалам наблюдения (фактор В — «интервал»: 6 уровней). Таким образом, фактор А — группа испытуемых животных — является *межгрупповым* фактором, поскольку он контролируется на различных группах испытуемых или на независимых выборках. Фактор В — *внутригрупповой* фактор, поскольку он контролируется внутри *одной и той же* группы крыс по отдельности.

Соотношение различных источников вариативности данных представлено на рис. 15 в виде схемы.

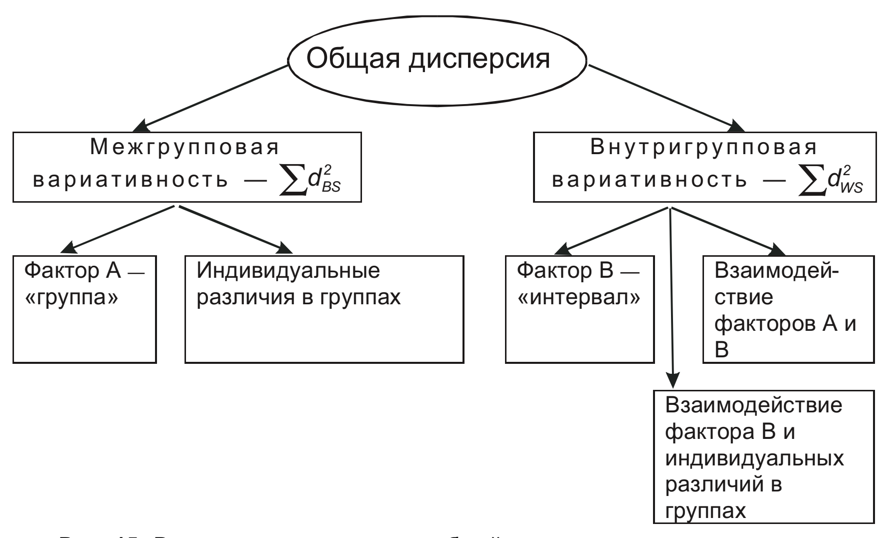


Рис. 15. Различные составляющие общей дисперсии в эксперименте со смешанным планом

[1] С некоторыми изменениями пример заимствован из книги Д.Хауэлла (1998).

Рассматривая вклад *межгрупповой вариативности* — $\sum d^2_{BS}$, мы выделяем в ней два компонента: 1) различия *между 3 группами* испытуемых (основной эффект фактора А) и 2) индивидуальные различия *внутри каждой группы* из 8 испытуемых. Первая компонента обусловлена включением в план эксперимента независимой переменной «группа испытуемых» (одна контрольная и две экспериментальные), что определяет такой источник вариативности как различия между группами испытуемых. Однако это не единственный источник различия между испытуемыми. Кроме того, в каждой отдельной группе есть 8 животных со своими индивидуальными различиями.

Фактор А в данном экспериментальном плане называют, таким образом, *межгрупповым фактором.* В принципе, если бы для каждой группы испытуемых усреднить данные по всем 6 интервалам наблюдения, то мы бы смогли провести обычный однофакторный ДА и оценить основной эффект фактора «группа». Проведя такой отдельный анализ, можно оценить также и соответствующую остаточную сумму квадратов ($\sum d^2_{BG}$), не зависимую от влияния (эффекта) индивидуальных различий.

Во *внутригрупповой вариативности* — $\sum d^2_{WS}$ мы выделяем три компонента (см. рис. 15): 1) различия между 6 последовательными *временными интервалами,* в рамках которых производилась оценка двигательной активности животных (т.е. основной эффект фактора В); 2) *взаимодействие* фактора В и фактора А и 3) *взаимодействие* фактора В с индивидуальными различиями испытуемых внутри групп. Первый компонент обусловлен введением в эксперимент 6 интервалов наблюдения за *одной и той же* группой крыс, поэтому фактор В и называют *внутригрупповым.* Он также зависит от индивидуальных различий *внутри* группы испытуемых. Поскольку фактор В — внутригрупповой, то и его взаимодействие с фактором А (второй компонент) также является внутригрупповым эффектом. Третий компонент внутригрупповой вариативности — взаимодействие фактора В и индивидуальных

различий внутри группы — иногда называют остаточным, поскольку, как мы отмечали выше, он фактически соответствует остаточной сумме квадратов[1].

В двухфакторном ДА со смешанным планом суммы квадратов связаны очень простыми соотношениями; их стоит (наряду с соответствующим рис. 15) внимательно изучить, чтобы более четко усвоить специфику соотношения межгрупповых и внутригрупповых факторов и лучше ориентироваться в стандартных распечатках большинства статистических систем:

$$\sum d^2_{total} = \sum d^2_{BS} + \sum d^2_{WS}, \qquad (43)$$

$$\sum d^2_{BS} = \sum d^2_{\text{Фактор } A} + \sum d^2_{IWG},, \qquad (44)$$

$$\sum d^2_{IWG} = \sum d^2_{\text{Фактор } B} + \sum d^2_{\text{Фактор } A \times B} + \sum d^2_{\text{Фактор } B \times IWG}. \qquad (45)$$

Аналогичными соотношениями связаны и соответствующие суммам квадратов величины степени свободы.

***Пример 2*: *два межгрупповых фактора и один внутригрупповой*[2].**

В исследовании Лауренса и др. (1995) изучалась эффективность нового вида обучающей программы (тренинга) по предупреждению риска венерических заболеваний среди чернокожих американцев. Использовались 2 типа обучающих тренингов: традиционный, двухчасовой и новый, восьмичасовой (фактор А — «тренинг»: 2 уровня). Участвовали испытуемые обоих полов — мужчины и женщины (фактор В — «пол»: 2 уровня). Зависимой переменной была частота использования презервативов в течение 4 периодов проведения исследования: до тренинга, сразу после него, спустя 6 и 12 месяцев (фактор С — «время»: 4 уровня). В исследовании приняли участие 4 группы испытуемых по 10 человек в каждой (гр. 1—гр.4). Таким образом, обрабатывались данные трехфакторного эксперимента 2×2×4. Данный экспериментальный план

---

[1] Фактически сумма квадратов взаимодействия фактора В и индивидуальных различий рассчитывается как сумма сумм квадратов отдельно *по каждой* группе. Поэтому этот компонент логически эквивалентен остаточной сумме квадратов, как мы ее рассматривали в предыдущих разделах.

[2] Пример с незначительными изменениями заимствован у Д.Хауэлла (1998).

представим в виде таблицы и разберем, какие факторы были межгрупповыми, а какой — внутригрупповым (см. табл. 33).

Таблица 33

Схема трехфакторного эксперимента с ДА с повторными измерениями: 2×2×4

	Традиционное обучение				Тренинг поведенческих навыков			
	До	После	6 мес.	12 мес.	До	После	6 мес.	12 мес.
Мужчины	гр. 1	гр. 1	гр. 1	гр. 1	гр. 2	гр. 2	гр. 2	гр. 2
Женщины	гр. 3	гр. 3	гр. 3	гр. 3	гр. 4	гр. 4	гр. 4	гр.4

Поскольку фактор «время» контролировался внутри каждой из четырех групп испытуемых и, тем самым, по каждой группе были 4 раза собраны *повторные измерения*, он является *внутригрупповым*. Таким образом, источником внутригрупповой вариации данных будут: 1) основной эффект фактора «время»; 2) взаимодействие основного эффекта фактора «время» и эффекта индивидуальных различий испытуемых в группе; 3) взаимодействие фактора «время» и фактора «тренинг»; 4) взаимодействие фактора «время» и фактора «пол»; 5) тройное взаимодействие факторов «время» × «тренинг» × «пол».

Факторы «группа» и «пол» — *межгрупповые*, поскольку их вариации наблюдаются на разных *независимых* выборках (группах) испытуемых. Источниками межгрупповой вариации являются: 1) основной эффект фактора «тренинг»; 2) основной эффект фактора «пол»; 3) взаимодействие факторов «тренинг» × «пол»; 4) эффект индивидуальных различий испытуемых.

Пример 3: два внутригрупповых фактора и один межгрупповой.

С заикающимися детьми проводилась психотерапевтическая работа с использованием двух альтернативных методов психотерапии (фактор «психотерапия»: 2 уровня). Психотерапевтическая работа сопровождалась медикаментозным лечением, включавшим 3 различные дозы приема миорелаксирующего препарата (фактор «доза»: 3 уровня). Качество речи

ребенка оценивалось в три временных периода исследования: до начала лечения, в середине лечения, в конце лечения (фактор «время»: 3 уровня). Схема эксперимента представлена в таблице 34.

Таблица 34

Схема трехфакторного эксперимента со связанными выборками: 3×2×3

	Психотерапия 1			Психотерапия 2		
	Доза 1	Доза 2	Доза 3	Доза 1	Доза 2	Доза 3
До лечения	гр. 1	гр. 1	гр. 1	гр. 2	гр. 2	гр. 2
В середине	гр. 1	гр. 1	гр. 1	гр. 2	гр. 2	гр. 2
В конце	гр. 1	гр. 1	гр. 1	гр. 2	гр. 2	гр. 2

Факторы «психотерапия» и «доза» — внутригрупповые, поскольку они контролируются на одной и той же группе испытуемых. Фактор «время» — межгрупповой, т.к. его эффект мы наблюдаем на двух независимых группах испытуемых.

Межгрупповая вариация состоит из двух компонентов: 1) основного эффекта фактора «время» и 2) эффекта индивидуальных различий испытуемых в каждой группе.

Внутригрупповая вариация более сложна и складывается из следующих составляющих: 1) основной эффект фактора «психотерапия»; 2) взаимодействие факторов «психотерапия»×«время»; 3) взаимодействие основного эффекта фактора «психотерапия» и эффекта индивидуальных различий испытуемых в группе; 4) основной эффект фактора «доза»; 5) взаимодействие факторов «доза»×«время»; 6) взаимодействие основного эффекта фактора «доза» и эффекта индивидуальных различий испытуемых в группе; 7) взаимодействие факторов «психотерапия»×«доза»; 8) тройное взаимодействие факторов «психотерапия»×«доза»×«время». Кроме того, в силу специфики взаимодействия двух внутригрупповых факторов появляется еще один источник вариации — *взаимодействие взаимодействия* этих факторов с эффектом индивидуальных различий испытуемых: («психотерапия»×«доза»)×«индивидуальные различия». Этот источник вариации часто называ-

ют *модельным*, поскольку, хотя он и оценивается, но какого-либо содержательного значения в анализе данных не играет.

Пример 4: только внутригрупповые факторы.

Такой экспериментальный план, когда все факторы — внутригрупповые, очень редко обсуждается в статистической литературе, хотя в психологических исследованиях встречается довольно часто. Например, так бывает при исследовании сенсорно-перцептивных процессов , когда необходимы уникальные испытуемые, найти которых чрезвычайно сложно, а поэтому, будучи однажды найденными, они долгое время участвуют в опытах[1].

В такого типа экспериментальном плане, когда в каждой клетке таблицы представлен всего один испытуемый, фактически каждая независимая переменная — это внутригрупповой фактор.

Линейную двухфакторную модель, лежащую в основе такого типа ДА, можно представить следующим образом:

$$X_{ij} = \mu_{общ.} + a_i + b_j + ab_{ij} + p_n + ap_{in} + bp_{jn} + abp_{ijn} + \varepsilon_{ijn} \qquad (46)$$

где: $\mu_{общ.}$ — общее среднее, a_i — дополнительный компонент, соответствующий влиянию *i*-го фактора, b_j — дополнительный компонент, соответствующий влиянию *j*-го фактора, ab_{ij} — компонент взаимодействия двух факторов, p_n — компонент, связанный с влиянием индивидуальности испытуемого, ap_{in} — компонент, обусловленный взаимодействием фактора *a* и эффектом индивидуальности испытуемого, bp_{jn} — компонент, обусловленный взаимодействием фактора *b* и эффектом индивидуальности испытуемого, abp_{ijn} — компонент, обусловленный тройным взаимодействием: факторов *a*, *b* и влиянием индивидуальности испытуемого, ε_{ijn} — ошибка линейной модели, связанная с *i*-м и *j*-м факторами и *n*-м испытуемым.

Хотя мы внесли в структурную модель составляющую, связанную со взаимодействием индивидуальности испытуе-

[1] Это могут быть испытуемые с уникальными сенсорными способностями, или, наоборот, с редкими дефектами зрения. Нередко исследования проводят в течение многих лет на одних и тех же парах близнецов.

мого и факторов, ее вклад, как правило, бывает очень незначителен и фактически плохо интерпретируем.

Рассмотрим в качестве примера вымышленное психогенетическое исследование, в котором в качестве испытуемых участвуют тройни (S_1, S_2 и S_3). В эксперименте оценивался индекс активации головного мозга по спектру электроэнцефалограммы (ЭЭГ). Первый контролируемый фактор — условие регистрации ЭЭГ: глаза открыты, глаза закрыты, сенсорная стимуляция (фактор «ЭЭГ-активация»: 3 уровня). Второй фактор — место регистрирующего электрода на поверхности головы: лоб — Fz, центр — Cz, темя — Pz, затылок — Oz[1] (фактор «зона»: 4 уровня). Третий фактор — возраст испытуемых: 1 год, 2 года, 3 года (фактор «возраст»: 3 уровня). Схема эксперимента представлена для наглядности в таблице 35.

Таблица 35

Схема эксперимента с тремя внутригрупповыми факторами: 3×4×3

Возраст	Глаза открыты				Глаза закрыты				Сенс. стимуляция			
	Fz	Cz	Pz	Oz	Fz	Cz	Pz	Oz	Fz	Cz	Pz	Oz
1 год	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3
2 года	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3
3 года	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3	S_1 S_2 S_3

Таким образом, трое одних и тех же испытуемых включены в каждую клетку таблицы, и, следовательно, они являются источником данных, полученных во всех 36 условиях эксперимента (3×4×3). Из этого следует, что все 3 фактора — внутригрупповые.

---

[1] Fz, Cz, Pz и Oz — стандартные обозначения мест расположения электродов на скальпе.

Рассмотрим составные части внутригрупповой вариации данных этого эксперимента: 1) основные эффекты всех трех факторов; 2) эффект индивидуальных различий испытуемых; 3) эффекты их попарного взаимодействия; 4) эффект тройного взаимодействия факторов; 5) эффекты взаимодействия индивидуальных различий испытуемых с парными взаимодействиями факторов; 6) эффект взаимодействия индивидуальных различий испытуемых с тройным взаимодействием факторов. Всего 15 составляющих. Как правило, для содержательной интерпретации интерес представляют лишь основные эффекты факторов и их взаимодействие.

Контрольные вопросы

1. Что означают понятия «независимые» и «связанные выборки» применительно к экспериментальному плану?

2. Покажите на примере, почему связанность выборок предполагает корреляцию между ними.

3. Какие преимущества имеет вариант ДА с повторными измерениями?

4. Что означает термин Repeated-Measures Design?

5. Каким образом в ДА с повторными измерениями компенсируется наличие зависимых (связанных) выборок?

6. Какой компонент общей дисперсии обозначается как $\sum d_{BS}^2$?

7. Напишите формулу структурной модели двухфакторного ДА с повторными измерениями и дайте ей пояснение.

8. Какое дополнительное математическое допущение принимается в ДА с повторными измерениями?

9. Что означают термины Within-Subjects Factor и Between-Subjects Factor?

10. Какие составляющие входят во внутригрупповую и межгрупповую вариативность?

4.6. Обработка данных с повторными измерениями в статистической системе STADIA 6.0

В системе Stadia имеется несколько вариантов обработки данных с повторными измерениями. Основой для программной реализации используемых алгоритмов ДА автору Stadia А.П.Кулаичеву послужила книга А.Афифи и С.Эйзена «Статистический анализ. Подход с использованием ЭВМ» (1982). К сожалению, как отмечают и сами авторы данной книги (с.266), в предложенных алгоритмах не учитывается факт зависимости повторных измерений. Это означает, что при оценке остаточной суммы квадратов мы *не исключили* влияние межиндивидуальной вариативности и тем самым ее скорректировали. Поэтому предложенные в системе Stadia методы обработки целесообразно использовать при обработке экспериментальных данных в том случае, когда в процессе исследования в какой-то степени обеспечена независимость выборок, полученных при повторных измерениях. Например, такую зависимость можно уменьшить, делая большие временные интервалы между последовательными исследованиями. Таким образом, имеющиеся в статистической системе Stadia алгоритмы ДА с повторными измерениями имеют ограниченное использование[1].

Четыре различных алгоритма ДА с повторными измерениями, реализованные в Stadia, соответствуют четырем различным типам экспериментальных планов, и поэтому при использовании данной статистической системы следует четко ориентироваться на описание этих моделей, данное в книге А.Афифи и С.Эйзена (с.264—266). Кратко опишем, как выполнить обработку данных для этих моделей. В меню статистических методов нужно выбрать подменю — *«Дисперсионный анализ»*, а в нем вариант — *«Двухфакторный»*. В появившемся окне даны 4 варианта моделей с повторными измерениями:

---

[1] В настоящее время А.П.Кулаичевым уже внесены необходимые изменения в алгоритм ДА.

1. Модель с *фиксированными эффектами* означает, что должны быть заданы 2 фактора с фиксированными эффектами.

2. Модель со *случайными эффектами* означает, что должны быть заданы 2 фактора со случайными эффектами.

3. Модель с *рандомизированными блоками* предполагает, что первый (внутригрупповой) фактор является фиксированным, а второй (межгрупповой) — случайным. Например, исследуется различие двух методов психотерапии (первый фактор) на группе из 5 испытуемых, причем каждый метод применяется к каждому испытуемому; в качестве измеряемой переменной используется какой-нибудь показатель самооценки. В данном случае блоком является испытуемый.

4. Модель *с группировкой* предполагает, что один из факторов — главный (группирующий), а второй — подчиненный (группируемый). Например, одна случайная выборка испытуемых подвергается одному методу психотерапии, а другая случайная выборка — другому. При этом с каждым испытуемым проводится по 5 повторных измерений. В данном случае метод психотерапии является группирующим фактором, по которому сгруппированы испытуемые (второй фактор). Первый фактор — фиксированный, второй — случайный.

При обработке данных указанных выше экспериментов исходные данные вводятся в *Редактор данных* особым образом. Они представляют собой псевдоматрицу (столбцы не должны быть обязательно одинаковой длины), в которой переменные (столбцы) должны соответствовать комбинации различных уровней исследуемых факторов, а строки — повторным измерениям. Причем столбцы должны следовать в порядке изменения уровней первого фактора: все уровни первого фактора для первого уровня второго фактора, все уровни первого фактора для второго уровня второго фактора и т.д. Сколько было измерено повторных сочетаний уровней двух факторов, столько и строк будет содержать матрица данных.

Результаты дисперсионного анализа включают стандартную таблицу, описанную выше в главе 3 (см. с.70).

4.7. Обработка данных с повторными измерениями в статистической системе SPSS 8.01

Как было сказано выше, в англоязычной литературе процедура ДА с повторными измерениями получила название *Repeated-Measures* (повторные измерения). В системе SPSS такая процедура имеется. В отличие от системы Stadia, в SPSS данный тип дисперсионного анализа организован как универсальная процедура, не зависящая от конкретного вида экспериментального плана, что позволяет психологу производить обработку данных широкого круга экспериментов, в которых могут быть выделены внутригрупповые факторы. Вместе с тем такая универсальность накладывает известные трудности, обусловленные необходимостью четкого выделения в проведенном исследовании внутригрупповых и межгрупповых факторов, ковариат и даже сознательного выбора нюансов модели ДА. Здесь, как и в других сложных статистических процедурах, за широту и универсальность приходится «платить» необходимостью знать и понимать различные варианты и тонкости ДА с повторными измерениями.

По сравнению со всеми предыдущими примерами, форма ввода данных в электронную таблицу для обработки с помощью данной процедуры значительно отличается. Данные заносятся следующим образом. Сначала вводят переменные, соответствующие уровням внутригруппового(ых) фактора(ов). Затем вводятся переменные, содержащие уровни межгруппового(ых) фактора(ов).

Рассмотрим, как обрабатываются данные двух примеров, описанных в данном разделе. Первый пример (см. табл. 30) представляет собой однофакторный план с повторными измерениями: на группе из 9 испытуемых, страдающих мигренью, в течение 5 недель изучалось влияние тренинга релаксации на частоту проявления головных болей. Внутригрупповым фактором в данном примере является время эксперимента: 5 уровней — 5 недель. Данные в *Редактор Данных* (SPSS Data

Editor) вводятся так, как это показано ниже во фрагменте электронной таблицы SPSS, т.е. как *последовательность столбцов*, соответствующих пяти уровням одного фактора:

	week1	week2	week3	week4	week5
1	21	22	8	6	6
2	20	19	10	4	4
3	17	15	5	4	5
4	25	30	13	12	17
5	30	27	13	8	6
6	19	27	8	7	4
7	26	16	5	2	5
8	17	18	8	1	5
9	26	24	14	8	9
10	.	.	.	.	.
11	.	.	.	.	.

Таким образом, девяти строкам соответствуют 9 испытуемых, пяти столбцам — 5 уровней фактора.

После окончания ввода данных в меню статистических методов *Статистика* (Statistics) следует выбрать группу методов *Общая линейная модель* (General Linear Model), а в ней — нужную нам процедуру ОЛМ — *Повторные измерения* (GLM — Repeated-Measures). Далее необходимо *определить*, с какими *внутригрупповыми* факторами (Within-Subject) мы будем работать, для чего в появившемся окне следует *указать имя* фактора. Можно задать *любое имя* — это никак не связано с именами введенных переменных. Определим имя нашего фактора: *Имя внутригруппового фактора* (Within-Subjects Factor Name): **неделя**. Далее следует указать число уровней этого фактора. Число уровней в нашем примере (Number of Levels) — **5**. Нажатием на кнопку *Добавить* (Add) мы вводим один внутригрупповой фактор в список факторов. Если для оценки изучаемого фактора в эксперименте использовалось несколько переменных (т.е. измерения про-

водились сразу по нескольким параметрам[1]), то можно указать сразу несколько переменных. Для этого следует нажать на кнопку *Измерения*>> (Measure>>), указать для каждой переменной (экспериментальной меры) подходящее *Имя измерения* (Measure Name:), а затем нажать на кнопку *Добавить* (Add). Поскольку в качестве единственной зависимой переменной в нашем примере измерено количество часов головной боли, дадим ей название «время». Окончание этапа определения факторов завершается нажатием на кнопку *Определить* (Define). В новом появившемся окне необходимо указать, какие *переменные* соответствуют уровням заданного фактора. Для этого выделяем щелчком мыши все 5 нужных нам переменных, и, нажав на кнопку со стрелочкой, переносим их в окно направо. Поскольку у нас нет ни межгрупповых факторов (Between-Subjects Factors), ни переменных-ковариат (Covariates), то можно сразу запустить процедуру ДА. Однако, в данной процедуре имеется несколько дополнительных возможностей, с которыми, на наш взгляд, полезно ознакомиться.

Пользователь может выбрать одну из *моделей обработки данных — Модель...* (Model...). Задание варианта *Полный факторный* (Full Factorial) означает, что рассчитываются все основные факторные эффекты (как внутригрупповых, так и межгрупповых факторов) и их взаимодействия. Вариант *Настраиваемая* (Custom) означает, что можно включить в модель только межгрупповые (Between-Subjects) и/или внутригрупповые (Within-Subjects) факторы, а также ограничить вычисления оценкой основных факторных эффектов или их взаимодействий. Изменять способ расчета *Сумм квадратов* (Sum of Squares:) по сравнению с предложенным по умолчанию (Type III) не очень опытным пользователям вряд ли стоит.

---

[1] Например, в нашем примере головная боль могла оцениваться не только по своей продолжительности, но также и по характеру проявления, локализации и т.д. Такие экспериментальные планы, где измерялась более чем одна независимая переменная, называют планами с многомерными повторными измерениями (Doubly Multivariate Repeated Measures Designs).

Если пользователь хочет оценить соотношение между уровнями фактора, то следует выбрать опцию *Контрасты* (Contrasts...), а в ней один из 6 методов оценки контрастов — по умолчанию контрасты не оцениваются.

При необходимости в результаты могут быть включены графики соотношения средних значений измеряемой переменной для различных уровней различных факторов. Для этого следует воспользоваться опцией *Графики* (Plots...)

Методы множественных парных сравнений реализуются в опции *Апостериори* (Post Hoc Tests) только для межгрупповых факторов, поэтому в нашем случае они не доступны.

Для того, чтобы рассчитанные параметры модели *сохранить* в матрице данных, предусмотрена возможность их записи как новых переменных — *Сохранить* (Save...). Эти новые переменные могут быть в дальнейшем проанализированы на предмет соответствия модели исходным данным.

В опции *Параметры* (Options...) можно заказать для выдачи в результаты анализа много дополнительных показателей, характеризующих данные и модель ДА.

Как и в предыдущих случаях, для начала выполнения данной процедуры ДА требуется нажать кнопку *OK*.

В окне результатов распечатывается довольно обширная статистическая информация, касающаяся не только оценки одного факторного эффекта, но также и описательной статистики, результатов теста на сферичность, проверки гомогенности дисперсий и т.д. Мы приведем здесь лишь основные результаты (см. табл. 36): дескриптивные статистики (табл. 36, *А*) и оценка эффекта внутригруппового фактора по многомерным критериям (табл. 36, *Б*) и *F*-критерию (табл. 36, *Г*), результаты теста на сферичность (табл. 36, *В*). Ограничимся лишь короткими комментариями полученных результатов. Как видно из распечатки результатов ДА, и многомерные тесты, и оценка по *F*-критерию исследуемого факторного эффекта (неделя наблюдения за головными болями) показали его высокую значимость. Это означает, что средние оценки времени головных болей, зарегистрированные испытуемыми после участия в тренинге релаксации, были различны. Обратив внимание на таблицу с результата-

Таблица 36

Результаты однофакторного ДА данных со связанными выборками в системе SPSS 8.01

Общая Линейная Модель

А

Within-Subject Factors

Измерение: ЧАСЫ

НЕДЕЛЯ	Зависимая переменная
1	НЕДЕЛЯ1
2	НЕДЕЛЯ2
3	НЕДЕЛЯ3
4	НЕДЕЛЯ4
5	НЕДЕЛЯ5

Descriptive Statistics

	Среднее	Стд. отклонение	N
НЕДЕЛЯ1	22.33	4.58	9
НЕДЕЛЯ2	22.00	5.34	9
НЕДЕЛЯ3	9.33	3.39	9
НЕДЕЛЯ4	5.78	3.42	9
НЕДЕЛЯ5	6.78	4.12	9

Б

Multivariate Tests

Эффект		Значение	F	Ст. св. гипотезы	Ст.св. ошибки	Знач.	Эта в квадрате
НЕДЕЛЯ	След Пиллая	.986	86.391	4.000	5.000	.000	.986
	Лямбда Уилкса	.014	86.391	4.000	5.000	.000	.986
	След Хотеллинга	69.113	86.391	4.000	5.000	.000	.986
	Наибольший корень Роя	69.113	86.391	4.000	5.000	.000	.986

В

Mauchly`s Test of Sphericity

Измерение: ЧАСЫ

Внутриобъектный эффект	W Моучли	Прибл. хи-квадрат	ст.св.	Знач.	Эпсилон		
					Гринхауз-Гайссер	Юнх-Фельдт	Ограниченный снизу
НЕДЕЛЯ	.282	8.114	9	.537	.684	1.000	.250

Проверка нулевой гипотезы о том, что ковариационная матрица ошибок ортонормированного преобразования зависимых переменных пропорциональна единичной матрице.

Г

Tests of Within-Subject Effect

Измерение: ЧАСЫ

Источник		Сумма квадратов типа III	ст.св.	Средний квадрат	F	Знач.	Эта в квадрате
НЕДЕЛЯ	Предполагая сферичность	2449.200	4	612.300	85.042	.000	.914
	Гринхауз-Гайссер	2449.200	2.738	894.577	85.042	.000	.914
	Юнх-Фельдт	2449.200	4.000	612.300	85.042	.000	.914
	Ограниченный снизу	2449.200	1.000	2449.200	85.042	.000	.914
Ошибка(НЕДЕЛЯ)	Предполагая сферичность	230.400	32	7.200			
	Гринхауз-Гайссер	230.400	21.903	10.519			
	Юнх-Фельдт	230.400	32.000	7.200			
	Ограниченный снизу	230.400	8.000	28.800			

ми оценки средних по неделям (см. табл. 36, *А* — дескриптивные статистики), можно сделать вывод о том, что интенсивность головных болей после тренинга в целом по группе испытуемых снизилась. Использование как многомерных, так и одномерного тестов было оправдано, поскольку тест сферичности Моучли показал (см. табл. 36, *В*), что у нас нет

достаточных оснований отвергать принятие этого предположения ($p = 0.537$).

Отметим, что в русской версии SPSS в качестве русских аналогов соответствующих английских терминов Between-Subjects Factor и Within-Subjects Factor использованы, на наш взгляд, не совсем удачные варианты перевода — внутриобъектные и межобъектные факторы (эффекты). Более того, в самой процедуре ОЛМ — *Повторные измерения* при определении этих факторов и в итоговых распечатках допущены технические ошибки, связанные с неправильным называнием факторов и переменных. Поэтому мы советуем читателю работать с англоязычной версией SPSS или пользоваться нашим описанием данной процедуры, ориентируясь на английские аналоги русских терминов, данные в скобках.

В следующем примере работы с SPSS обработаем данные эксперимента, в котором изучалась эффективность нового вида обучающей программы (тренинга) по предупреждению риска венерических заболеваний среди чернокожих американцев (см. с.96). Напомним, что экспериментальный план был трехфакторным: *два межгрупповых фактора и один внутригрупповой*. В качестве возможного варианта работы с SPSS покажем читателю, как пользоваться англоязычной версией этой системы.

Как и в предыдущем примере, сначала вводятся данные в соответствии с внутригрупповым фактором — Time (время исследования). Поскольку наблюдение за испытуемыми велось в 4 временных периода, то в первые 4 переменные (pretest, posttest, fu6, fu12) последовательно заносим соответствующие данные. Каждый столбец-переменная включает 40 наблюдений: (10 мужчин + 10 женщин) × 2 вида тренинга. В пятую переменную (gender) вносим коды пола испытуемых: 1— мужчины, 2 — женщины. В шестую переменную (training) заносим данные о виде тренинга, в котором участвовали испытуемые: 1 — традиционное обучение, 2 — тренинг поведенческих навыков.

После ввода данных и проверки его правильности переходим в меню статистических методов (Statistics), где

	pretest	posttest	fu6	fu12	gender	training
1	7.00	22.00	13.00	14.00	1.00	2.00
2	25.00	10.00	17.00	24.00	1.00	2.00
3	50.00	36.00	49.00	23.00	1.00	2.00
4	16.00	38.00	34.00	24.00	1.00	2.00
5	33.00	26.00	24.00	26.00	1.00	2.00
6	10.00	7.00	23.00	26.00	1.00	2.00
7	13.00	33.00	27.00	24.00	1.00	2.00
8	22.00	20.00	21.00	11.00	1.00	2.00
9	4.00	.00	12.00	.00	1.00	2.00
10	17.00	16.00	20.00	18.00	1.00	2.00
11	.00	6.00	22.00	26.00	2.00	2.00
12	.00	18.00	12.00	15.00	2.00	2.00
13	.00	8.00	.00	.00	2.00	2.00

Рис. 16. Ввод данных в электронную таблицу SPSS

выбираем General Linear Model (*Общая линейная модель*), а ней — процедуру Repeated-Measures (*Модель с повторными измерениями*). Далее *определяем* внутригрупповой (Within-Subject) фактор, для чего в появившемся окне *указываем его имя* — **time** (время) и число уровней фактора: Number of Levels: **4**. Нажимаем на кнопку *Add* (добавить) и вводим наш один внутригрупповой фактор в список факторов. Далее указываем название измеряемого параметра независимой переменной, для чего, нажав на кнопку *Measure*>>(мера, параметр), пишем имя (Measure Name:) — **freq** (сокращение от слова *частота*) и вносим его в список, нажимая на кнопку *Add*. Заканчиваем этап определения факторов нажатием на кнопку *Define*. В новом появившемся окне указываем, какие *переменные* соответствуют уровням заданного фактора: *последовательно* выбираем переменные pretest, posttest, fu6, fu12 и переносим их в правое окно.

Поскольку у нас имеются *два межгрупповых фактора* (Between-Subjects Factors), указываем, каким переменным они соответствуют: gender и trening, и переносим их в окно

Таблица 37

Результаты ДА с одним внутригрупповым и двумя межгрупповыми факторами в системе SPSS 8.0

General Linear Model

Within-Subjects Factors

Measure: FREQ

TIME	Dependent Variable
1	PRETEST
2	POSTTEST
3	FU6
4	FU12

Between-Subjects Factors

		N
GENDER	1.00	20
	2.00	20
TRENING	1.00	20
	2.00	20

Box's Test of Equality of Covariance Matrices

Box's M	45.811
F	1.210
df1	30
df2	3563
Sig.	.200

Tests of Within-Subjects Effects

Measure: FREQ

Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
TIME	293.819	3	97.940	.936	.426	2.807	.250
TIME * GENDER	832.419	3	277.473	2.651	.052	7.953	.634
TIME * TRENING	1465.319	3	488.440	4.667	.004	14.000	.884
TIME * GENDER * TRENING	528.669	3	176.223	1.684	.175	5.051	.430
Error(TIME)	11304.025	108	104.667				

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: FREQ

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Intercept	33785.156	1	33785.156	68.781	.000	68.781	1.000
GENDER	3267.056	1	3267.056	6.651	.014	6.651	.709
TRENING	124.256	1	124.256	.253	.618	.253	.078
GENDER * TRENING	77.006	1	77.006	.157	.694	.157	.067
Error	17683.275	36	491.202				

a. Computed using alpha = .05

Between-Subjects Factor(s).

В качестве *модели обработки данных* (Model...) выбираем вариант Full Factorial (*Полная факторная модель*), чтобы получить оценки всех основных факторных эффектов (как внутригруппового, так и межгрупповых факторов) и их взаимодействий. В окошке *Расчет суммы квадратов* (Sum of Squares:) оставляем предложенный по умолчанию тип 3 (Type III).

Оценку контрастов (Contrasts...) и графики (Plots...) не заказываем.

Для двух межгрупповых факторов расчет парных сравнений (Post Hoc Tests) заказывать не следует, поскольку тесты парных сравнений выполняются в SPSS в том случае, если число уровней факторов превышает 2.

Сохранять параметры модели в матрице данных (Save...) не будем.

В *Опциях* (Options...) заказываем тест на гомогенность дисперсий выборок — ставим «галочку» в окошке *Gomogeneity Tests*.

Следует помнить, что все сделанные выборы в каждом из окошек нужно подтверждать нажатием на кнопку *Continue*.

Начало выполнения ДА инициируем нажатием кнопки *OK*.

Рассмотрим некоторые результаты проведенного ДА (см. табл. 37). В верхней части таблицы приводится резюме по всем факторам и результаты теста Бокса на симметричность ковариационной матрицы. Тест Бокса не обнаружил отклонений от симметричности ($p = 0.2$). Далее в распечатке представлены результаты анализа влияния *внутригруппового* фактора Time. Его основной эффект оказался незначим ($F = 0.936$; $p = 0.426$), тогда как эффект его взаимодействия с межгрупповым фактором Training статистически высоко достоверен ($F = 4.667$; $p = 0.004$) Эффект взаимодействия со вторым межгрупповым фактором Gender оказался на грани достоверности ($F = 2.651$; $p = 0.052$). Эффект тройного взаимодействия факторов оценен как недостоверный ($F = 1.684$; $p = 0.175$).

Ниже даны результаты анализа межгрупповых факторов. Значимым оказался основной эффект фактора Gender ($F = 6.651$; $p = 0.014$), а *основной эффект фактора Training*, и *взаимодей-*

ствие этих факторов статистически не значимы.

Задания

1. Дайте содержательную интерпретацию результатов ДА первого и второго примера, обработанных в системе SPSS.

2. В группе из 8 студентов, изучающих второй иностранный язык на четвертом курсе, в течение 3 недель последовательно проводились тесты на знание английского языка (как основного иностранного). В таблице, приведенной ниже, дано число ошибок, допущенных каждым

Студент	Этапы проверки знаний		
	1 неделя	2 неделя	3 неделя
1	55	57	58
2	44	44	47
3	61	63	61
4	65	67	67
5	40	46	45
6	70	68	71
7	49	51	51
8	58	55	59

из студентов в серии из 100 заданий:

Напишите структурную модель для ДА этих данных. Проведите ДА с помощью одной из статистических систем. Дайте содержательную интерпретацию полученных данных.

3. Проводилось исследование влияния эффективности обратной связи, даваемой испытуемым экспериментатором на успешность обучения сенсомоторным навыкам. В эксперименте участвовали 2 группы по 10 испытуемых — экспериментальная (давалась обратная связь) и контрольная (не было обратной связи). Ниже представлены данные об успешности деятельности испытуемых в двух сериях эксперимента — фоновой (до начала тренировки) и тестовой (после оконча-

Испытуемый	Экспериментальная группа		Контрольная группа	
	фоновая	тестовая	фоновая	тестовая
1	8	9	3	5
2	5	7	5	5
3	3	2	8	10
4	5	7	2	5
5	2	9	5	3
6	6	7	6	10
7	5	8	6	9
8	6	5	4	5
9	4	7	3	7
10	4	9	5	6

ния тренировки):

Проведите ДА, используя соответствующую процедуру в любой статистической системе. Дайте содержательную интерпретацию полученных результатов. Постройте графики соотношения средних значений по факторам.

4. В исследовании факторов[1], способствующих курению, опрашивались испытуемые, бросившие курить тремя различными способами: постепенное уменьшение выкуренных сигарет, резкое прекращение курения и использование аверсивной терапии. В каждой группе из 5 испытуемых предлагалось оценить по 10-балльной шкале желание испытуемых «покурить сейчас же». Испытуемые опрашивались в двух различных ситуациях — дома и на работе. Опрос проходил в 2 этапа — до прекращения курения и

Как бросил курить	Время опроса			
	до того, как бросил курить		после того, как бросил курить	
	Место курения			
	дома	на работе	дома	на работе
Постепенно	7	6	6	4
	5	4	5	2
	8	7	7	4
	8	8	6	5
	6	5	5	3
	8	7	7	6
	5	5	5	4
Резко	7	6	6	5
	8	7	6	5
	7	6	5	4
	9	8	5	4
	4	4	3	2
С терапией	7	7	5	3
	7	5	5	0
	8	7	6	3

после.

В этом эксперименте у нас один межгрупповой фактор — «Способ бросить курить» и два внутригрупповых

[1] Пример заимствован из книги Д.Хауэлла (1998).

фактора — «Место курения» и «Время опроса». Проведите соответствующий ДА и проинтерпретируйте полученные результаты.

5. Д.Хауэлл пытался доказать своей дочери, что постоянное использование карманного калькулятора в школе и дома отнюдь не способствует развитию навыков устного счета. Чтобы быть более убедительным, он провел небольшое исследование. Среди одноклассников дочери он выбрал 5 человек, у которых был калькулятор, и 5 — у которых его не было. Каждый ребенок прошел 3 устных арифметических теста на сложение, вычитание и умножение. Были получены следующие оценки успешности устных вычислений по

Наличие калькулятора	Сложение	Вычитание	Умножение
есть	8	5	3
	7	5	2
	9	7	3
	6	3	1
	8	5	1
нет	10	7	6
	7	6	5
	6	5	5
	9	7	8
	9	6	9

10-балльной шкале:

Проведите ДА и установите, подтвердилась ли гипотеза Д.Хауэлла о «вреде» калькуляторов.

Глава 5

Многомерный дисперсионный анализ — MANOVA

5.1. Специфика многомерного ДА

Как было отмечено выше, для модели ДА с повторными измерениями (Repeated-Measures Analysis of Variance) важно выполнение допущения о симметричности ковариационной матрицы или в более общей форме — сферичности. Хотя точно оценить степень влияния нарушения этого допущения на результаты ДА достаточно сложно, тем не менее рядом авторов было показано, что ДА со связанными выборками чувствителен к нарушению сферичности при оценке контрастов и простых эффектов.

Поэтому многие авторы предлагают использовать другую процедуру ДА — *многомерный дисперсионный анализ*. В англоязычной литературе она получила название MANOVA (Multivariate Analysis of Variance). Данная процедура ДА не требует выполнения допущения о сферичности. Тем не менее, многомерный ДА — это более сложная процедура и в ряде случаев менее мощная с точки зрения статистического вывода, особенно на выборках небольшого размера. Поскольку в психологических исследованиях использование малых выборок скорее правило, чем исключение, то при использовании многомерного ДА психологу следует обратить на это особое внимание.

Как следует из названия, *многомерный ДА* — это процедура ДА, имеющая дело с *несколькими зависимыми переменными* одновременно. В психологических экспериментах часто вводится несколько зависимых переменных. Например, в психологии личности нередко используют данные, полученные

по 3—5 шкалам многофакторных личностных опросников, в педагогической психологии измеряют успеваемость, мотивацию, социометрический статус учеников, в психофизиологии при измерении параметров вызванного потенциала измеряют амплитуды и латентности нескольких его компонентов и т.д. Таким образом, каждый испытуемый получает сразу по несколько измерений, которые подвергаются ДА одновременно, в противоположность тому, чтобы последовательно выполнить ДА для каждой переменной в отдельности.

Казалось бы, зачем использовать такую дополнительную процедуру ДА, когда можно несколько раз последовательно провести одномерный ДА с каждой из зависимых переменных? Это на первый взгляд резонное предложение не учитывает *влияние интеркорреляции между самими зависимыми переменными*. При игнорировании связи между зависимыми переменными мы лишаемся важного источника информации о структуре полученных данных. Таким образом, процедура многомерного ДА основана на более общей и более универсальной модели экспериментальных данных по сравнению с одномерным ДА. Одномерный ДА является частным случаем многомерного.

С другой стороны, если у исследователя нет оснований считать, что используемые им зависимые переменные связаны корреляционной зависимостью, то, по-видимому, нет и достаточных оснований для применения многомерного ДА, а следует ограничиться одномерным. Для оценки корреляционной связи зависимых переменных можно исследовать корреляционную матрицу этих переменных: если они независимы, то недиагональные элементы матрицы будут незначительно отличаться от нуля. *Тест Бартлетта на сферичность* обычно используется для строгой статистической оценки корреляционной матрицы: он проверяет, является ли корреляционная матрица *единичной матрицей*, т.е. матрицей с диагональными элементами, равными 1, и нулевыми недиагональными элементами. Этот тест основан на вычислении *детерминанта* матрицы ошибок корреляционной матрицы: если детерминант близок к нулю, то одна или не-

сколько зависимых переменных могут быть выражены как линейные функции других зависимых переменных. Таким образом, статистическая *гипотеза о независимости* используемых в эксперименте переменных *отвергается*, если детерминант не отличен от нуля, и, следовательно, корреляционная матрица не является единичной.

5.2. Допущения многомерного ДА

Поскольку многомерный ДА имеет дело с несколькими переменными, то делается допущение о характере их *совместного распределения*: в отличие от одномерного ДА, здесь требуется, чтобы переменные имели *многомерное нормальное распределение*. Второе допущение многомерного ДА требует, чтобы для каждого уровня фактора переменные образовывали одну и ту же дисперсионно-ковариационную матрицу. Как следует из названия, дисперсионно-ковариационная матрица — это квадратная матрица, на диагонали которой лежат дисперсии переменных, а ее элементами являются коэффициенты ковариации между переменными.

5.3. Многомерные критерии для оценки факторных эффектов

В силу наличия в структурной модели многомерного ДА нескольких зависимых переменных и эффекта их корреляции, для многомерного ДА были разработаны отличные от привычных для одномерного ДА сумм квадратов статистические показатели вариативности. Это так называемые *многомерные тесты*. Эти показатели основываются на вычислении детерминанта произведения двух матриц — матрицы, называемой *матрицей сумм квадратов и коэффициентов ковариации* (H) на обратную *матрицу остаточных сумм квадратов и коэффициентов ковариации* (E^{-1}) — $H \times E^{-1}$. Как следует из их названия, диагональные элементы матриц представляют собой, соответственно, суммы квадратов или остатки сумм квадратов по каждой зависимой переменной, а недиагональные элементы — коэффициенты ковариации между ними. Как известно детерминант матрицы является мерой общей вариабельности данных или дисперсии

всей матрицы. Для тех читателей, кто не знаком с матричной алгеброй и факторным анализом, напомним, что детерминант может быть вычислен как произведение собственных значений матрицы, поскольку каждое собственное значение является, в свою очередь, частью общей дисперсии. Фактически вычисление собственных значений матрицы $H \times E^{-1}$ реализуется посредством ее факторизации методом главных компонент.

Обычно в компьютерных статистических программах вычисляются несколько статистических показателей для многомерной оценки различий между сравниваемыми средними:

1. Критерий Пиллая (Pillai's trace):

$$V = \sum_{i=1}^{S} \frac{1}{1+\lambda_i}\,; \qquad (47)$$

2. Критерий Уилкса (Wilks's lambda):

$$W = \prod_{i=1}^{S} \frac{1}{1+\lambda_i}\,; \qquad (48)$$

3. Критерий Хотеллинга (Hotelling trace):

$$T = \sum \lambda_i\,; \qquad (49)$$

4. Критерий Роя (Roy's largest root):

$$R = \frac{\lambda_{MAX}}{1+\lambda_{MAX}}\,; \qquad (50)$$

где λ_i — i-ое собственное значение матрицы, λ_{MAX} — максимальное собственное значение матрицы, S — количество ненулевых собственных значений матрицы.

Для каждого из приведенных выше многомерных критериев построены собственные распределения, однако все они могут быть преобразованы в статистики, хорошо аппроксимирующиеся F-распределением. При одной независимой переменной все 4 указанных выше критерия эквивалентны обычной F-статистике одномерного ДА.

При выборе наилучшего из 4 критериев обычно исходят из двух важных соображений — *мощность* критерия и его *устойчивость* к нарушению допущений ДА (или *робастность*). Для большинства практических ситуаций использования мно-

гомерного ДА мощность критериев убывает согласно их месту в следующем списке: Пиллая, Уилкса, Хотеллинга и Роя. Критерий Пиллая является также самым робастным критерием. Это означает, что уровень значимости этого критерия оценивается вполне корректно даже в том случае, когда нарушаются допущения ДА. Последнее немаловажно, поскольку многие исследователи не затрудняют себя выполнением статистических проверок математических допущений ДА.

Контрольные вопросы

1. Каковы основные отличия многомерного ДА от одномерного?
2. Можно ли сказать, что многомерный ДА по своим математическим допущениям более либеральная процедура?
3. Почему важно убедиться, что между включенными в многомерный ДА зависимыми переменными имеется корреляционная связь?
4. Почему результаты нескольких одномерных ДА в принципе неэквивалентны многомерному ДА?
5. Для чего используется тест Бартлетта на сферичность?
6. Какие вы знаете многомерные статистические критерии, оценивающие факторный эффект, и чем они отличаются от одномерного F-критерия?
7. Какой из 4 многомерных критериев считается самым мощным и устойчивым к нарушениям допущений многомерного ДА?

5.4. Пример обработки данных многомерным ДА

В силу сложности вычислительного алгоритма многомерного ДА мы опустим конкретные вычисления факторных эффектов и их взаимодействий и ограничимся лишь самыми необходимыми пояснениями при разборе одного примера и его обработки. Поскольку во многих статистических системах (BMDP, SAS, SPSS) имеются полезные предустановки параметров модели многомерного ДА (установки по умол-

чанию), даже не очень искушенному в статистике исследователю использование этой процедуры становится доступным.

Обратимся вновь к данным Д.А.Кинга, проводившего эксперимент по изучению привыкания крыс к воздействию наркотика (1986). Кинг регистрировал двигательную активность животных сразу же после *инъекции мидозалама*. После первой дозы наблюдалось заметное снижение двигательной активности. Но, как и все морфины, мидозалам вызывал быстрое привыкание. Кинг исследовал динамику такого привыкания. Использовались три *различные группы* по 8 животных (фактор А — «группа»: 3 уровня) — контрольная и две экспериментальных. Экспериментальные группы отличались тем, что первая группа крыс («старые условия») исследовалась в тех же окружающих условиях, в которых она и содержались, тогда как условия для второй группы были изменены («новые условия»). Контрольная группа крыс получала наркотик впервые только в день эксперимента, а две экспериментальные группы получали инъекции в течение нескольких дней до контрольного тестирования. Поскольку использованный наркотик оказывает быстрое действие, особенности двигательного поведения животных в каждой группе оценивались в течение часа после его приема по 6 пятиминутным интервалам наблюдения. Таким образом, в контрольной группе исследовалась *первичная реакция* на наркотик, а в двух экспериментальных группах — типичный *эффект привыкания*.

Поскольку автора главным образом интересовал эффект привыкания, то, выполняя многомерный ДА, мы будем оценивать динамику двигательной активности во времени для трех групп животных. 5 переменных будут являться мерами изменения моторной активности животных во времени. Экспериментальный план Д.А.Кинга выглядит следующим образом: 5 зависимых переменных и один межгрупповой фиксированный фактор (см. табл. 38).

Выполняя многомерный ДА, для каждой переменной мы проверяем *нулевые гипотезы* о равенстве средних показателей двигательной активности в трех различных группах животных.

Таблица 38

Данные эксперимента Д.А.Кинга

Группа крыс	Интервал наблюдения двигательной активности				
	инт.1	инт. 2	инт. 3	инт. 4	инт. 5
контрольная	150	71	59	132	74
	335	156	160	117	230
	149	91	115	43	154
	159	127	212	71	224
	159	35	75	71	34
	292	184	246	225	170
	297	66	96	209	74
	170	42	66	114	81
в старых условиях	346	177	192	239	140
	426	236	76	102	232
	359	183	123	183	30
	272	82	85	101	98
	200	263	216	241	227
	366	263	164	220	180
	371	270	308	219	267
	447	294	216	284	225
в новых условиях	282	225	134	189	169
	317	85	120	131	205
	362	144	114	115	127
	338	91	77	108	169
	263	141	142	120	195
	138	16	95	39	55
	329	62	6	93	67
	292	104	184	193	122

Рассмотрим, как обрабатываются данные нашего примера в системе SPSS. Результаты измерений вводятся в *Редактор Данных* (Data Editor) в виде шести переменных, соответствующих шести интервалам наблюдения: в каждой переменной по 8×3=24 испытуемых, т.е. данные по каждой группе животных вводятся в столбец-переменную последовательно. Во фрагменте электронной таблицы SPSS (см. рис. 17) они обозначены как *инт1, инт2 ... инт6*. Седьмая переменная (*группа*) кодирует принадлежность животного к одной из экспериментальных групп: 1 — контрольная группа, 2 — живот-

	int1	int2	int3	int4	int5	int6	group	var	var
1	150	44	71	59	132	74	1.00		
2	336	270	156	160	117	230	1.00		
3	149	52	91	115	43	154	1.00		
4	159	31	127	212	71	224	1.00		
5	159	0	35	75	71	34	1.00		
6	292	126	168	246	226	178	1.00		
7	297	107	66	96	209	74	1.00		
8	170	37	42	66	114	91	1.00		
9	346	175	177	192	239	140	2.00		
10	426	329	236	76	102	232	2.00		
11	359	238	183	123	183	38	2.00		
12	272	60	82	65	101	98	2.00		
13	200	271	263	216	241	227	2.00		

Рис. 17. Пример ввода данных эксперимента Д.А.Кинга в *Редактор Данных* SPSS

ные в старых условиях, 3 — животные в новых условиях. Таким образом, в итоге мы получаем матрицу данных 24×7.

После окончания ввода данных в меню статистических методов *Статистика* (Statistics) следует выбрать группу методов *Общая линейная модель* (General Linear Model), а в ней — нужную нам процедуру ОЛМ — *Многооткликовая* (Multivariate...). Далее необходимо *указать зависимые переменные,* для чего выделяем их в левом окне и, нажимая на кнопку со стрелочкой, переносим в окно *Зависимые переменные* (Dependent Variables). Затем следует указать, какая переменная соответствует фактору. Также выделяем ее в левом окне и переносим в окно *Фиксированные факторы* (Fixed Factors), т.е. факторы с фиксированными уровнями. Поскольку в нашем примере нет переменных-ковариат (Covariates), то в соответствующем окне мы ничего не указываем.

В процедуре многомерного ДА при работе с данными появляется еще одна возможность: числовые значения каждой переменной могут быть трансформированы посредством процедуры взвешивания по методу наименьших квадратов: *Взвешенный МНК* (WLS Weight — weighted least-squares). Эта воз-

можность заключается в том, что в данных можно указать отдельную переменную, с помощью которой результаты проведенных измерений получат различный вес. Подобная трансформация, как правило, выполняется при необходимости компенсировать различия в точности измерений используемых в анализе переменных. Пользователям, не знакомым с использованием данного приема, мы не рекомендуем его применение.

Пользователь может выбрать одну из *моделей обработки данных* ДА: *Модель* (Model...). Выбор варианта *Полная факторная* (Full Factorial) означает, что рассчитываются все основные факторные эффекты (как внутригрупповых, так и межгрупповых факторов) и их взаимодействия. Вариант *Настраиваемая* (Custom) означает, что можно выбрать оценку основных факторных эффектов или их взаимодействия, а также и ограничить уровень факторного взаимодействия. Поскольку в нашем примере имеется лишь один фактор, то мы выбираем вариант *Полная факторная*.

Об изменении способа расчета сумм квадратов (опция *Сумма квадратов* — Sum of Squares:) стоит думать в случае наличия более, чем одного фактора и неравного количества испытуемых в группах (так называемые несбалансированные данные). Только в этом случае появляется неоднозначность в решении задачи разделения общей дисперсии на составляющие. Это связано с тем, что оцениваемые различия между средними зависят от влияния эффектов других факторов. Поэтому для проверки разных гипотез можно использовать различные методы расчета сумм квадратов. Например, предложенный по умолчанию метод (Тип III) соответствует так называемому *уникальному*, *или регрессионному*, алгоритму, когда при оценке любого факторного эффекта учитываются результаты оценки *всех остальных* факторных эффектов. Для того, чтобы разобраться в подобных тонкостях, которые лишь незначительно влияют на результат, мы советуем читателю обратиться к руководству по использованию конкретной статистической системы. Предложенный по умолчанию *Тип III* (Type III) является наиболее универсальным выбором. В окне выбора модели предлагается еще одна возможность, специ-

фичная для многомерного ДА — оценить отличие от нуля сравниваемых средних по различным группам и переменным: *Включить в модель свободный член* (Include intercept in model). В различных статистических системах эта функция может называться по-разному: MEAN, CONSTANT или INTERCEPT. Эта функция может быть очень полезна для психологов в том случае, если включенные в анализ переменные представляют собой отклонения индивидуальных оценок испытуемых от тестовой нормы. Тогда как дополнительный результат анализа мы получим по каждому фактору и каждой переменной статистическую оценку отличия данной выборки испытуемых от нормы. Это сравнение выполняется также с помощью многомерных тестов.

Если пользователь хочет оценить соотношение между уровнями фактора, то следует выбрать один из 6 методов оценки контрастов: *Контрасты* (Contrasts...), по умолчанию контрасты *не оцениваются*. Опытным пользователям при необходимости работы с контрастами мы советуем обязательно обратиться к соответствующему описанию статистической системы SPSS. Неопытным коллегам рекомендуем пользоваться установками по умолчанию.

При необходимости в результаты могут быть включены графики соотношения средних значений измеряемой переменной для различных уровней различных факторов: *Графики* (Plots...).

Среди методов множественных парных сравнений (*Апостериори* — Post Hoc Tests) выберем тест Геймса—Хауэлла.

Для того, чтобы рассчитанные параметры модели *сохранить,* в матрице данных предусмотрена возможность их записи как новых переменных: *Сохранить* (Save...). Эти новые переменные могут быть предметом специального анализа на предмет соответствия обработанных данных используемой модели ДА.

Дополнительно можно заказать для выдачи в результаты анализа много дополнительных показателей, характеризующих данные и модель ДА: *Параметры* (Options...)

Как и в предыдущих случаях, для начала выполнения данной процедуры ДА требуется нажать кнопку *OK*.

В окне результатов может быть представлена очень обширная информация, состоящая из нескольких страниц таблиц и графиков. Мы приведем здесь лишь основные результаты (см. табл. 39) и сделаем необходимые комментарии.

В верхней части таблицы (см. табл. 39, *А*) представлены результаты проверки допущений многомерного ДА. М-тест Бокса оценивает равенство дисперсионно-ковариационных матриц для каждого уровня фактора. Полученные результаты показывают ($F = 0.929$, $p = 0.603$), что у нас нет оснований говорить о нарушении этого допущения. Результаты теста Бартлетта на сферичность показывают, что корреляционная матрица переменных не является единичной матрицей, следовательно наши переменные связаны корреляционной связью (аппрокс.$\chi^2 = 73.873$, $p < 0.005$).

В следующей части таблицы (см. табл. 39, *Б*) даны результаты всех четырех отмеченных выше многомерных тестов: по всем тестам основной эффект фактора «группа животных» (Group) оказался высоко значимым. Таким образом, проверяемая нами нулевая гипотеза о равенстве групповых средних может быть отвергнута. Еще ниже представлены результаты уже известных нам тестов одномерного ДА по каждой из 6 переменных. Видно, что по каждой переменной обнаружен статистически достоверный эффект фактора «группа животных» ($p < 0.005$). Для того, чтобы разобрать, какие средние различаются между собой, в самой нижней части таблицы (см. табл. 39, *Г*) представлены результаты множественных (6 сравнений по каждой из 6 переменных) парных сравнений по критерию Геймса—Хауэлла.

В статистической системе SPSS имеется возможность включать в многомерный ДА не только межгрупповые, но и внутригрупповые факторы, однако эта возможность реализуется с помощью специального командного режима выполнения статистических процедур (Comand Syntax). С этими дополнительными возможностями можно ознакомиться в специальном руководстве по использованию Comand Syntax, либо обратившись к режиму «Помощь» (Help) из основного меню и выбрав в нем раздел Syntax Guide, а в нем — подраздел Advanced Statistics.

Таблица 39

Результаты многомерного ДА в системе SPSS 8.01

Общая Линейная Модель

А

Between-Subjects Factor

		N
группа животных	1.00	8
	2.00	8
	3.00	8

Box's Test of Equality of Covariance Matrices

Box's M	69.121
F	.929
df1	42
df2	1309
Sig.	.603

Bartlett's Test of Sphericity

Likelihood Ratio	.000
Approx. Chi-Square	73.873
df	20
Sig.	.000

Б

Multivariate Tests

Эффект		Значение	F	Ст. св. гипотезы	Ст.св. ошибки	Знач.
Intercept	След Пиллая	.960	64.105	6.000	16.000	.000
	Лямбда Уилкса	.040	64.105	6.000	16.000	.000
	След Хотеллинга	24.039	64.105	6.000	16.000	.000
	Наибольший корень Роя	24.039	64.105	6.000	16.000	.000
ГРУППА	След Пиллая	.841	2.057	12.000	34.000	.049
	Лямбда Уилкса	.308	2.140	12.000	32.000	.043
	След Хотеллинга	1.765	2.207	12.000	30.000	.039
	Наибольший корень Роя	1.425	4.038	6.000	17.000	.011

В

Tests of Between-Subjects Effects

Источник	Зависимая переменная	Сумма квадратов типа III	ст.св.	Средний квадрат	F	Знач.
Скорректированная модель	1 интервал	72793.000	2	36396.500	6.260	.007
	2 интервал	152425.333	2	76212.667	9.719	.001
	3 интервал	75468.000	2	37734.000	9.645	.001
	4 интервал	16913.083	2	8456.542	1.851	.182
	5 интервал	30403.583	2	15201.792	4.089	.032
	6 интервал	9037.000	2	4518.500	.895	.424
Intercept	1 интервал	1937448.4	1	1937448.4	333.218	.000
	2 интервал	555104.167	1	555104.167	70.792	.000
	3 интервал	483936.000	1	483936.000	123.693	.000
	4 интервал	448540.042	1	448540.042	98.182	.000
	5 интервал	527770.042	1	527770.042	141.955	.000
	6 интервал	524808.375	1	524808.375	103.986	.000
ГРУППА	1 интервал	72793.000	2	36396.500	6.260	.007
	2 интервал	152425.333	2	76212.667	9.719	.001
	3 интервал	75468.000	2	37734.000	9.645	.001
	4 интервал	16913.083	2	8456.542	1.851	.182
	5 интервал	30403.583	2	15201.792	4.089	.032
	6 интервал	9037.000	2	4518.500	.895	.424
Ошибка	1 интервал	122101.625	21	5814.363		
	2 интервал	164668.500	21	7841.357		
	3 интервал	82160.000	21	3912.381		
	4 интервал	95937.875	21	4568.470		
	5 интервал	78075.375	21	3717.875		
	6 интервал	105985.625	21	5046.935		
Итог	1 интервал	2132343.0	24			
	2 интервал	872198.000	24			
	3 интервал	641564.000	24			
	4 интервал	561391.000	24			
	5 интервал	636249.000	24			
	6 интервал	639831.000	24			
Скорректированный итог	1 интервал	194894.625	23			
	2 интервал	317093.833	23			
	3 интервал	157628.000	23			
	4 интервал	112850.958	23			
	5 интервал	108478.958	23			
	6 интервал	115022.625	23			

Таблица 39 (продолжение)

Г

Множественные сравнения

Геймс—Хауэлл

Зависимая переменная	(I) группа животных	(J) группа животных	(I-J)-я разность средних	Стд. ошибка	Знач.	95% доверительный интервал Нижняя граница	Верхняя граница
1 интервал	1.00	2.00	-134.50*	38.126	.012	-238.53	-30.47
		3.00	-76.25	38.126	.138	-173.85	21.35
	2.00	1.00	134.50*	38.126	.012	30.47	238.53
		3.00	58.25	38.126	.296	-39.76	156.26
	3.00	1.00	76.25	38.126	.138	-21.35	173.85
		2.00	-58.25	38.126	.296	-156.26	39.76
2 интервал	1.00	2.00	-171.50*	44.276	.012	-304.90	-38.10
		3.00	-5.00	44.276	.991	-107.46	97.46
	2.00	1.00	171.50*	44.276	.012	38.10	304.90
		3.00	166.50*	44.276	.008	48.81	284.19
	3.00	1.00	5.00	44.276	.991	-97.46	107.46
		2.00	-166.50*	44.276	.008	-284.19	-48.81
3 интервал	1.00	2.00	-124.50*	31.275	.004	-206.87	-42.13
		3.00	-12.00	31.275	.912	-88.79	64.79
	2.00	1.00	124.50*	31.275	.004	42.13	206.87
		3.00	112.50*	31.275	.012	25.53	199.47
	3.00	1.00	12.00	31.275	.912	-64.79	88.79
		2.00	-112.50*	31.275	.012	-199.47	-25.53
4 интервал	1.00	2.00	-43.88	33.795	.480	-140.79	53.04
		3.00	19.63	33.795	.805	-62.34	101.59
	2.00	1.00	43.88	33.795	.480	-53.04	140.79
		3.00	63.50	33.795	.175	-24.49	151.49
	3.00	1.00	-19.63	33.795	.805	-101.59	62.34
		2.00	-63.50	33.795	.175	-151.49	24.49
5 интервал	1.00	2.00	-75.88	30.487	.088	-161.91	10.16
		3.00	-.75	30.487	1.000	-77.46	75.96
	2.00	1.00	75.88	30.487	.088	-10.16	161.91
		3.00	75.13	30.487	.058	-2.37	152.62
	3.00	1.00	.75	30.487	1.000	-75.96	77.46
		2.00	-75.13	30.487	.058	-152.62	2.37
6 интервал	1.00	2.00	-44.75	35.521	.498	-146.19	56.69
		3.00	-8.50	35.521	.964	-95.55	78.55
	2.00	1.00	44.75	35.521	.498	-56.69	146.19
		3.00	36.25	35.521	.562	-55.58	128.08
	3.00	1.00	8.50	35.521	.964	-78.55	95.55
		2.00	-36.25	35.521	.562	-128.08	55.58

*. Разность средних значима на уровне .05.

Задания

1. Двум группам из 6 испытуемых-добровольцев в каждой вводились две различные дозы стимулирующего лекарства (фактор «доза»: 2 уровня). В ходе эксперимента измерялось их функциональное состояние по поведенческому (порог слуховой чувствительности) и физиологическому (частота пульса — ЧСС) показателям. В таблице приведены экспериментальные данные в условных единицах относительно исходного (фонового) уровня функционального состояния, оцененного до эксперимента.

Обработайте полученные данные методом многомерного ДА в одной из статистических систем и оцените влияние дозы лекарства на каждую зависимую переменную.

Первая группа испытуемых			Вторая группа испытуемых		
Доза	Порог	ЧСС	Доза	Порог	ЧСС
1	3.4	2.2	2	3.3	2.8
1	3.4	2.2	2	3.2	2.6
1	3.3	2.3	2	3.2	2.7
1	3.4	2.3	2	3.2	2.6
1	3.3	2.2	2	3.2	2.7
1	3.3	2.0	2	3.3	2.6

2. В экспериментах А.Н.Гусева и С.А.Шапкина (неопубликованные данные, 1994) изучалось влияние времени суток (межгрупповой фактор 1: 2 уровня — «утро» и «вечер») и мотивации достижения испытуемых (межгрупповой фактор 2: 2 уровня — «мотивированные на избегание неудачи» и «мотивированные на успех») на эффективность обнаружения зрительных сигналов в условиях повышенной бдительности. В таблице приведены значения 4-х показателей времени реакции испытуемых (ВР).

Обработайте полученные данные методом многомерного ДА в одной из статистических систем и оцените влияние времени суток и мотивации достижения на каждую зависимую переменную: среднее ВР на правильные обнаружения — ВР(Hit), среднее ВР на ложные тревоги — ВР(FA), среднеквадратичное отклонение ВР на правиль-

ные обнаружения — $o_{BH(Hit)}$ и среднеквадратичное отклонение ВР на ложные тревоги — $o_{BH(FA)}$. Оцените для каждой переменной эффект взаимодействия факторов «время суток» и «мотивация достижения».

Есть ли различие в оценках факторных эффектов и их взаимодействия, проведенных с помощью многомерных тестов (то, что делает MANOVA) и *F*-критерия (то, что делает ANOVA)?

№ исп-го	ВР(Hit)	ВР(FA)	$O_{BH(Hit)}$	$O_{BH(FA)}$	Время дня	Мотивация
1	285	345	97	143	2	1
2	257	268	36	103	2	1
3	289	324	38	40	1	2
4	266	238	93	95	2	2
5	264	258	30	38	2	2
6	370	378	77	107	1	2
7	269	252	50	45	2	1
8	331	298	88	64	1	1
9	356	300	124	68	1	1
10	416	355	71	51	1	2
11	352	331	75	77	2	2
12	283	232	51	41	2	1
13	373	324	81	74	1	1
14	305	252	49	36	1	1
15	298	356	29	49	1	2
16	361	359	70	88	2	2
17	367	314	35	41	1	2
18	368	266	111	48	2	2
19	347	354	72	85	1	2
20	325	285	73	53	2	2
21	285	287	34	31	1	2

3. В исследовании Л.Я.Дорфмана и Г.Я.Ковалевой (неопубликованные данные, 1999) изучалась связь различных индивидуальных особенностей испытуемых. Каждый испытуемый оценивался по выраженности четырех субмодальностей Я: «Я — авторское», «Я — воплощенное»,

«Я — превращенное», «Я — вторящее» (Я1, Я2, Я3 и Я4)[1]. Кроме того у каждого испытуемого оценивались 2 показателя когнитивного стиля «концептуальная дифференциация» (КД1 и КД2). По степени оригинальности мышления (креативности) испытуемые были разбиты на 2 группы — с высокими и более низкими показателями оригинальности (межгрупповой фактор: 2 уровня). Авторов данного исследования интересовал вопрос о связи показателей субмодальностей Я и указанного выше когнитивного стиля с уровнем оригинальности (креативности) испытуемых.

Обработайте приведенные ниже в таблице данные методом многомерного ДА в одной из статистических систем и оцените влияние оригинальности на 6 указанных выше зависимых переменных.

Есть ли различие в оценках эффекта фактора «оригинальность», проведенных с помощью многомерных тестов и *F*-критерия?

№ исп-го	Я1	Я2	Я3	Я4	КД1	КД2	Оригинальность
1	42	43	54	38	7	10	1
2	52	48	49	35	12	9	1
3	41	45	45	31	11	6	2
4	40	45	40	48	14	7	1
5	48	46	51	42	6	15	2
6	44	42	45	21	17	9	1
7	35	45	43	38	7	12	2
8	51	50	53	28	10	10	1
9	39	47	54	36	7	8	2
10	40	44	49	36	12	7	1
11	44	41	49	35	33	12	2
12	47	49	50	34	9	10	1
13	43	48	48	37	11	7	1
14	49	47	51	38	17	8	2
15	44	47	53	39	9	8	2
16	48	41	46	26	9	7	2

[1] Четыре субмодальности Я рассматриваются автором в рамках его психологической концепции метаиндивидуального мира как некие сущностные характеристики Я (см. *Дорфман Л.Я.*, 1993—1998).

Приложение

Критические значения F-распределения для уровня значимости p = 0.05

Число ст. св. знаменателя	Число степеней свободы числителя									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	161.4	199.5	215.8	224.8	230.0	233.8	236.5	238.6	240.1	242.1
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
50	4.03	3.18	2.79	2.56	2.40	2.22	2.20	2.13	2.07	2.03
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85
1000	3.85	3.01	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84

Рекомендуемая литература

Афифи А., Эйзен С. Статистический анализ. Подход с использованием ЭВМ. М.: Мир, 1982.

Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. М.: Прогресс, 1976.

Готтсданкер Р. Основы психологического эксперимента. М.: Изд-во Моск. ун-та, 1982.

Кимбл Г. Как правильно пользоваться статистикой. М.: Финансы и статистика, 1982.

Корнилова Т.В. Введение в психологический эксперимент. М.: Изд-во Моск. ун-та; ЧеРо, 1997.

Кулаичев А.П. Методы и средства анализа данных в среде Windows. Stadia 6.0. М.: Российское психологическое общество, 1998.

Лакин Г.Ф. Биометрия. М.: Высшая школа, 1980.

Налимов В.В. Теория эксперимента. М.: Наука, 1971.

Руководство пользователя. SPSS Base 8.0. М.: СПСС Русь, 1998.

Руководство по применению. SPSS Base 8.0. М.: СПСС Русь, 1998.

Сидоренко Е.В. Методы математической обработки в психологии. СПб.: Социально-психологический центр, 1996.

Суходольский Г.В. Основы математической статистики для психологов. 2-е изд. СПб.: Изд-во СПб. ун-та, 1998.

Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере. М.: Инфра-М, 1998.

Шапкин С.А. Экспериментальное изучение волевых процессов. М.: Смысл, 1997.

Шеффе Г. Дисперсионный анализ. М.: Наука, 1980.

Dancey C., Reidy J. Statistics Without Maths for Psychology. Usung SPSS for Windows. Prentice Hall, 1999.

Howell D. Fundamental Statistics for the Behavioral Sciences (3rd ed.). Belmont, CA: Duxbury Press, 1995.

Howell D. Statistical Methods for Psychology (4th ed.). Belmont, CA: Duxbury Press, 1998.

Levine G. A Guide to SPSS for Analisis of Variance. Lawrence Erlbaum Associates, Publishers. Hillsdale, 1991.

Kirk R. Experimental Design: Procedures for the Behavioral Sciences. (3rd ed.) Pacific Grove, CA: Brooks/Cole, 1995.

Norusis M. SPSS Inc. SPSS for Windows: Advanced Statistics, Release 6.0. Chicago: SPSS Inc, 1993.

SPSS Advanced Statistics 7.5. Chicago: SPSS Inc, 1997.

Содержание

Глава 4

Дисперсионный анализ данных с повторными измерениями

Глава 5

Многомерный дисперсионный анализ — MANOVA

Алексей Николаевич Гусев

Дисперсионный анализ в экспериментальной психологии

Учебное пособие

Редактор М.И.Черкасская
Компьютерная верстка и дизайн обложки А.И.Чекалиной

Учебно-методический коллектор «ПСИХОЛОГИЯ»
Лицензия № 00451 от 15 ноября 1999 г.

Адрес: 107005, г. Москва, ул. Бауманская, д. 50/12, стр. 1.
103009, г. Москва, ул. Б. Никитская, д. 4.
Тел. (095)203-35-71, тел./факс (095)203-35-65.
E-mail: collect@mail.ru.

Подписано в печать 26.06.2000. Формат 60×84/16. Тираж 1000.
Бумага офсетная. Гарнитура TimesET. Усл. печ.л. 8,5